Theory of Spatial Pricing and Market Areas

Theory of Spatial Pricing and Market Areas

M. L. Greenhut and H. Ohta

Duke University Press Durham, N.C. 1975

L.C.C. card no. 74–83786
I.S.B.N. no. 8223–033–7

printed in the United States
by Kingspport Press

Abstract

Related to the firm's pricing pattern over economic space is the question of the size and shape of the market areas over which firms sell. This question applies either to the spatial monopoly market or to spatial competitors. It will be shown that their price and output practices are functionally related to the particular size and shape of the market areas over which they sell. In particular, it will be demonstrated that the monopolist sells over a circular-shaped market area while spatial competitors sell over hexagonal-shaped market areas. Most important, these results will be seen to apply regardless of whether or not the sellers discriminate in the pricing of their goods. Perhaps most surprising in the perspective of classical economic theory is our thesis that spatial competitors will discriminate often in their pricing over space. The economic factor which determines whether or not they will *discriminate* is the ratio of fixed costs to the buyers' limit price for the good.

Contents

Preface

This book is the outgrowth of individual as well as collaborative research. It stems in part from early studies of pricing and location in economic space, and from research dealing with the interaction of spatial prices and market areas. Its focal point is our belief that the f.o.b. mill and discriminatory price practices of "spatial" monopolists are not contained at all in the monopoly price models of classical "spaceless" economic theory. In fact, it will be claimed that spatial and spaceless competitors differ even more sharply in their pricing practices than do spatial and spaceless monopolists. Many competitors who sell over economic space will be revealed to maximize profits by discriminating in price, even when the intensity of the competition eventuates in zero profits.

A major difference—indeed *the* major difference—between spaceless and spatial price theory lies, therefore, in the premise that in classical economic theory competitors were conceived only to price nondiscriminatorily while spatial competitors will be claimed generally to discriminate in price. Related to this difference is the thesis that the pricing policies of spatial sellers are inextricably tied to the sizes and shapes of their market areas. Certain sizes (as determined in part by freight rates, production costs, and conditions of entry) lead to discrimination, with only very special sizes of market areas requiring f.o.b. mill pricing. It is, furthermore, contended herein that spatial extensions of classical theory have been ignored for too long a period of time, and that the full tie-up between monopolistic and competitive pricing over economic distance must be considered and included in microeconomic theory.

It warrants mention that the subjects and theory set forth in this book could have been presented in the context of location theory or, more generally, the theory of the firm in economic space. However, these facets of the study of spatial prices are not included directly herein. Single emphasis is devoted instead to pricing techniques and their effects on the outputs and profits of firms and the distances and market shapes over which they sell. This "singular" emphasis reflects our belief that the present subject is an integral one in its own right. It also reflects the belief that readers who would want to relate the theory of spatial prices directly to such interest matters as plant location theory or the efficiency of the space economy will find enough direction in the present study for these or other extensions of our models.

We are indebted to several colleagues and friends at Texas A&M University. We wish, specifically, to thank Ray Battalio, Alfred Chalk, Robert Ekelund, Erik Furubotn, David Maxwell, David Nelson, Glen Williams, and the late Charles Ferguson for suggestions, advice, and editorial and research assistance. To Mrs. Eleanor Brown and Rosemary Bates go special thanks for their painstaking secretarial and typing efforts. An "in-house" research council grant to the department of economics provided initial computer time and other resources so helpful in undertakings of the present order, and we thank John Allen, David Maxwell, and Andrew Suttle in this regard. Not only did a National Science Foundation grant facilitate this and related empirical research now in process, but even more precisely, encouragement by James Blackman prompted us to set forth our theory of spatial prices and market areas in the form of this book. We would like also to thank Donald Dewey for suggestions related to Chapter 5, Jack Scheidell for permission to use co-authored material published previously, John Greenhut and Ming Hwang for ideas, critiques, and use of selected materials formed by each. We also acknowledge a debt to the *Manchester School*, the *Journal of Regional Studies* (Japan), the *American Economic Review*, *Weltwirtschaftliches Archiv*, the *Western Economic Journal*, *Environment and Planning* for use in revised form, chiefly in two appendixes of this book and in Chapters 2, 4, 5, 9, 11, of portions of materials published in those journals by one or both of the authors. We must finally thank Michael Proctor for permitting re-use in Chapter 10 of co-authored portions of a paper presented at the Second World Congress of the Econometric Society in London, September 1970. Without the help and permissions indicated here, this book would have been a much less enjoyable experience than it actually has been.

M. L. G.
H. O.

Theory of spatial pricing and market areas

Part 1. Regional (spatial) economics

1. Facets of economic space

I. *Introduction*

The subject of economic space has been acquiring new dimensions in recent years as classical location theory no longer remains the single (or even main) field of study dealing with the effect of distance on economic theory. Instead, economic space is regarded today as a multifaceted subject. Thus present-day analysis of the urban economy extends beyond the classic plant layout, site selection inquiries of yesteryear. Even the location of public facilities, the macroeconomic relations between the new suburbs and the old downtown districts, the length of time in travel to work, in particular the difficulties in getting to work of the urban poor as opportunities increase in the suburbs, are all part of the new urban-space economics. Such traditional subjects as minimum wage laws, seniority rules, unemployment of workers, fiscal policy, and inflation have acquired a spatial character and role as the urban landscape has come within the purview of the economist. The questions of decay of central business districts (CBD), the circular or sectoral form of the economic and residential development of cities, or, more generally, the location of economic activity in and outside of cities no longer hold a central place in the research of urban-space economists.

This is not to say that interest in the market shapes of firms, of the trading (or interest) areas of public facilities, and of the cities themselves has waned.[1] The so-called range of a good[2] which in another semantic helps delimit the maximum size of the hexagonal trading area of the city's CBD,[3] and the "threshold" of the good, which in the other semantic relates to the minimum size of the hexagonal scopes of its shopping centers,[4] are still vital topics today. Present-day extensions to these aspects of classical location theory have been far-reaching. But these extensions have transgressed the subject of plant location to include insights into

1. Interest in urban economics and particularly the trading areas of cities has clearly been increasing. Indeed, an important side inquiry in a book by one of the authors was dedicated hopefully towards the end of producing a "theory of the trading area of the city along the lines of the theory of the firm." See Greenhut [5, chaps. 10–13].
2. See Reilly [10] and Sweet [11].
3. See Christaller [1].
4. See Lösch [7].

such questions as whether increases in the sizes of city populations and urban areas are associated with increases or decreases in the concentration of people; hence, they also relate to (the problems of) congestion, energy requirements, and pollution in urban places. It is simply the case that old subjects and inquiries have been acquiring expanded or new roles.

One need have only an elementary appreciation of the concept of opportunity cost in economic theory to appreciate its relevance to the space economy. The matter of residential patterns, city growth and decay, and, more generally, development of the regional landscape all have roots in the (lost) opportunities of people. Thus the attraction of and interactions between the residents of cities and their institutions are specifiable in terms of gravity and potential models. But these models are only elementary reflections of the opportunities open to decision makers in economic space. In related manner, present-day emphasis on probabilities in the physical sciences has a counterpart use in studies dealing with social relations, an emphasis which is clearly evidenced in the *new* (more advanced) *gravity models* [5] of regional economics. The focal points of space economics thus extend all the way from opportunity costs to such forces as the economic base of the city, thence to population growth, congestion, energy requirements, and pollution. To cast a large picture narrowly but pointedly, the externalities which have long disturbed welfare economists may appear directly in *(and be resolved directly by)* the probabilistic gravity models under conception today.[6]

There is no need in this introduction to consider the still "broader" aspects of economic space which have appeared in the form of *interregional* input-output tables, *interregional* programming models, and the like. Any recountal on this score would carry us more and more to the field of aggregative economics and regional economic development, and indeed even back to comparative advantage and international finance. The interest of the present book lies chiefly (in fact, exclusively) along microeconomic lines, and it accordingly behooves us to restrict our introductory chapter to this theme. A *generalized* statement in broad description of the microeconomics of economic space will help provide the desired focus for some specific remarks entered later in this chapter about the contents of this book.

II. *Welfare economics and space microeconomics*

The firm in economic space differs from its nonspatial counterpart in diverse ways. For one thing, the very postulate of distance as a cost places

5. See Dacey [2, pp. 5–7].
6. See Olsson [9].

the theorist outside the fundamental context of the purely competitive framework, unless one assumes all buyers are at a point so that only sellers are differentiated by the varying distances between them and the buyers.[7] Indeed, it has been shown that in any nontrivial assumption of distances (i.e., where both individual buyers and sellers are distributed over the space), an oligopolistic market arises, to which behavioral uncertainty is intrinsic.[8] This uncertainty itself requires a new format (or matrix if you will) in decision making and in the locating of plants and cities, to say nothing about explaining regional growth and the like. Behavioral uncertainty might, in fact, appear to eliminate the determinacy of the marketplace, though the contrary has been contended elsewhere by one of the present authors.[9] It is significant, however, that whatever the welfare effects and properties may be of the firm in economic space, the practices of spatial sellers must be derived and evaluated. The pricing patterns of a firm or firms at a given point in space and those of competitors distributed over the economic landscape tend to differ in theory from the pricing patterns of spaceless firms. Sellers not subject to cost of distance, and even competitive sellers scattered over a costly space whose customers are all located at a single point in space, price differently than the spatial seller studied herein.

The requirement of one price in the market, or relevance of a lowest price possible to be charged a single buying point in space, gives way to ranges (and sets) of prices under nontrivial conceptions of competition over economic space. The net price of a good may be higher for a seller on sales to selected points in the space under discriminatory pricing and lower on sales to other points in the space, or his mill (net) price may be the same on all sales, with only delivered prices being different, etc. The subject of special interest to us in this book, spatial price discrimination, actually possesses features which lead to theorems running counter to those of nonspatial price discrimination theory. The quantity of output, the profitability of the firm, the distances over which the firm sells, the maximum and minimum sizes of the firm's market area, even the shapes of these market areas—all these change as a discriminatory pricing system is substituted for f.o.b. mill pricing. The changes involve and reflect differences in basic conclusions of spatial microeconomic theory compared with nonspatial microeconomic theory.

Even the antitrust aspects of price discrimination differ as one views a space economy in place of a nonspace economy. But given this observation, it deserves emphasis that the subject matter of this book is confined to unorganized (competitive) oligopolists;[10] hence, impacts of spatial

7. See Greenhut [4].
8. See Greenhut [3, chap. 2, sec. 3].
9. See Greenhut [5].
10. The term organized oligopoly is used to refer to oligopolists who pursue price leadership practices, or are in collusion via trade associations, gentlemen's agreements, cartel

economics on proposals to make antitrust policy more effective are basically *non sequitur* to this book.[11] Perhaps all that is required in these introductory remarks is the statement that our theory of spatial price discrimination diverges from nonspatial price discrimination theory in two important ways: (i) with respect to output effects and (ii) the "naturalness" of the policy. This second difference relates to the condition that in certain cases *competitive* firms in economic space tend naturally to discriminate in price; manifestly, views of effective antitrust policy could change as a result of this perspective.

An important difference existing between spatial and nonspatial price discrimination theory involves determination of the demand and cost conditions which generate (or promote) discriminatory pricing in place of nondiscriminatory pricing. Such determination is central to practically every chapter which follows.

III. *Plan of study*

To facilitate reading, we have relegated selected mathematical derivations – indicated by roman numeral superscripts – to the end of the book. On the other hand, the appendixes consist of materials closely related to the contents of the chapters of which they form parts. Accordingly, all mathematical formulations considered vital to a full understanding of chapter materials are included in the chapter.

Our practices with respect to figure and equation references will be evident in the text. But a caveat is in order about algebraic notation: to simplify printing, and in the process to reduce production costs as much as possible, we have followed standard journal recommendations dealing with mathematical notation. Thus we have used primes and asterisks, superscripts and subscripts rather than bars and circumflexes above letters. In one respect, this practice caused us to run out of sufficient number of notational alternatives; accordingly, a single prime superscript will be used herein to indicate (as is customary) the first derivative of a function *and in selected chapters,* namely, only 7–11, it is used whenever the equation relates to a state of competition. Most important, the meaning intended by this single prime will be clear from (or stressed in) the text.

---

organizations, and their like. The Pittsburgh-plus system (or basing-point system) is an example of organized oligopoly pricing in economic space. By unorganized oligopoly is meant competitive oligopoly, i.e., the competitors of Joan Robinson's imperfectly competitive market who identify each other. We shall see that oligopolists may follow either a discriminatory pricing system over economic space or a nondiscriminatory f.o.b. mill price system (where typically the term f.o.b. price is used herein to imply nondiscriminatory pricing over economic space).

11. The locational effects of spatial price discrimination are, also, not analyzed herein, except obliquely in Appendix II to Chapter 10 and in Chapter 11.

Chapters 2–6 consider the nonspatial as well as a spatial view of monopoly pricing. They examine the similarities and differences between nondiscriminatory and discriminatory monopoly pricing in both systems of thought. They also examine the differences in profits, prices, outputs, and sales distances of the spatial monopolist, given the alternatives of f.o.b. mill pricing and spatial price discrimination.

These and later chapters assume specific functions in some of their models, e.g., negatively sloping linear demand, zero or constant marginal cost of production, linear distribution of buyers, etc. These assumptions are made chiefly to simplify the argument, though occasionally they also serve as the basis for operational models which make experimental analysis possible (e.g., Chapter 6). Most important, the conclusions to be derived will be demonstrated elsewhere in the book to hold independently of them.

The assumption sets used in Chapters 2–6 – in fact throughout the book – contain theoretically important hypotheses. In particular, it is assumed (1) that freight costs are significant, (2) that for stress on the purely spatial properties of the subject, buyers are evenly or *homogeneously* distributed over the space, and (3) that each buyer's demand function is *identical* to that of any other buyer. The significance of the first assumption should be manifest. Without this assumption *economic* space cannot be evaluated: costless space is not economic space. The second assumption becomes significant only in connection with the first assumption. If freight costs are nil, assumptions (2) and (3) would yield the obvious conclusion that there can (and will) be no price discrimination. In contrast, when assumption (2) is combined with assumptions (1) and (3), it will be shown that the necessary implication is for sellers to discriminate in price. Then the impact of economic space on the pricing practices of the firm is unambiguously identifiable. In a similar pattern, whether spatially oriented price discrimination yields more (or less) outputs than nondiscriminatory pricing over space is determinable only by abstracting from the factors other than economic space which would influence the magnitudes of the demand functions. Assumptions (1), (2), and (3) are therefore intrinsic to the subject at hand. More specialized statements as to the contents of Chapters 2–6 are provided below. Readers seeking only a broad preview of our plan of study may omit all indented remarks.

> Chapter 2 compares spaceless and spatial demand and sets forth the fundamental principles of simple monopoly pricing (i.e., the nonspatial theory of monopoly pricing) vis-à-vis its counterpart in economic space, f.o.b. mill pricing.[12] The freight absorption applicable to profit-maximizing f.o.b. mill prices is also presented in this

12. Cf. the definitions given in note 10 of organized and unorganized oligopoly.

chapter. Finally, suggestions as to the relationship between spatial prices and market areas are briefly entered.

Chapter 3 turns to the subject of price discrimination, both in the classical spaceless and spatial systems of thought. It reviews the literature on the subject, and in particular evaluates the widespread belief that the shapes of demand curves determine whether or not a monopolist prices discriminatorily. The importance claimed for the shapes of demand curves is crystallized in the well-known (non-spatially oriented) theorem which holds that price discrimination may occur if and only if the elasticity of demand differs in distinctively identifiable markets.[13] Economists subscribing to this theory typically emphasize the relative shapes of the demand curves as the sole determinant of the question whether discriminatory price practice yields more outputs than does nondiscriminatory pricing. The spatial economists considered in Chapter 3 will, however, be seen to ignore this facet of the subject completely.

Chapter 4 proves that spatial price discrimination yields more outputs than does f.o.b. pricing, regardless of the "basic character (or shape) of the demand curve." This result stems from the fact that in the eyes of the seller the net demand curves of differentially spaced buyers *shrink* successively as the distances to the buyers increase. In contrast, classical nonspatial price discrimination theory dealt essentially with two (not innumerably many) separate markets, sales to each of which either involved zero freight cost or the same freight cost. The problem for the spatial monopolist is then not simply to compare (the) two separate markets, but to identify (i) the distant submarket(s) which is (are) to be included in the sales area, and (ii) the direction of his discrimination over space.

Chapter 5 establishes the proposition that the spatial monopolist necessarily discriminates in favor of more distant buyers regardless of the "basic character (or shape) of the demand curve." This proposition *requires* the proviso that the demand curve must be subject to the fundamental *economic* constraint that demand vanishes at some finite price. This proviso is one of the chief points of distinction between our spatial theory and the classical nonspatial theory of price discrimination.

Chapter 6 reexamines the two fundamental propositions set forth in Chapters 4 and 5 by means of a special "spatial" model which assumes three differently shaped demand curves, namely, a concave, a linear, and a convex demand. The model used in Chapter 6 enables one to view the alternative price schedules, distances, and outputs applicable to an economic landscape given (i) the different demand possibilities, and (ii) the alternative pricing practices of the firm (i.e.,

13. For example, see Henderson and Quandt [6, pp. 170–72].

spatial price discrimination or f.o.b. mill pricing). The claim that spatial price discrimination is efficient in the sense of providing greater outputs at generally lower price schedules marks the underlying thesis of this chapter.

Chapters 7 and 8 deal with competitive oligopoly, both in the spaceless and spatial form. These chapters compare and contrast the classical theoretical derivation of oligopoly pricing with its spatial variant. They indicate that price discrimination may be practiced by spatial competitors as well as by the spatial monopolist.

Chapter 7, in particular, compares nonspatial competitive oligopoly price theory with the counterpart f.o.b. mill price of the competitive spatial oligopolist. It is contended here that just as the nonspatial monopoly price has a direct transform in f.o.b. mill monopoly pricing, the spatially competitive oligopolist derives his profit-maximizing f.o.b. mill price in a way corresponding to that of the spaceless competitive oligopolist. At the same moment, the points of difference with respect to monopoly outputs and demand curves that were set forth in Chapters 4–6 will be seen to apply to competitive (spaceless and spatial) oligopolists.

Chapter 8 then compares the profits, price schedules, outputs, and distances of competitive oligopolists who practice spatial price discrimination with the profits, price schedules, outputs, and distances of competitive oligopolists who pursue f.o.b. mill pricing. This examination in terms of spatial competition is a counterpart to that of Chapter 6. Price schedules and related findings are derived for the competitive space economy which correspond to those relevant to a spatial monopoly. However, a competitive oligopoly market has long-run as well as short-run features. The effect of free (continuing) entry on the firm's price policy and outputs must also be determined. The model followed in Chapter 8 indicates that the price pattern offering the greatest distances and the lowest schedule of delivered prices ultimately depends on the extent of the competition prevailing in the space. This competition, in turn, is found to depend partly on the prevailing production and transportation cost levels. In particular, when production and transportation costs are small, discriminatory pricing gives way to nondiscriminatory pricing, *as the intensity of the spatial competition increases.* Whether the firm practices discriminatory or nondiscriminatory pricing depends ultimately, in other words, on the relative importance of costs (freight as well as production costs) in the zero-profit competitive equilibrium.[14] This chapter thus derives the "switching point" between the

14. The word "relative" relates to the highest price consumers would pay for a unit of the good.

alternative pricing techniques studied, in addition to presenting the competitively reduced profit-maximizing f.o.b. mill and discriminatory price schedules over an economic space.

Chapters 9, 10, and 11 conclude the book by examining the interdependence between *pricing techniques, final prices, market areas* (sizes and shapes), *profits,* and *free entry or exit of firms.* The theory of spatial market shapes that was formed chiefly by Lösch [7] and altered by Mills and Lav [8] will be reexamined along with related derivations of profit-maximizing prices. In some ways, the price derivations of Chapter 2 are carried forward here not only from the perspective of a spatial monopolist who is subject to a specific shape of market area, but from the specific perspectives of spatial oligopolists who are in short-run *and* long-run situations and are able to determine (or have determined for them) the desired (or requisite) size and shape of the market area.

Chapter 9 reestablishes a classical proposition on market area configurations that was first set forth by Lösch and subsequently criticized by Mills and Lav. It shows that spatial competitors necessarily sell in hexagon-shaped market areas except for what is revealed to be a trivial case. This chapter also formulates the precise relations between prices, sizes, and shapes of trading areas under conditions of spatial monopoly and spatial competition.

Chapters 10 and 11 conclude the book by comparing respectively the market sizes and shapes of competitive nondiscriminating oligopolists in short-run and long-run equilibria with sellers who discriminate over the space. The type of price and distance schedules which provide windfalls or fail to yield windfalls are determined in the context of market shapes relevant to a spatially competitive oligopolistic economy.

Though the locational impacts and welfare aspects of spatial price discrimination compared with f.o.b. mill pricing are not examined in this book,[15] it should be implicitly (if not intuitively) clear that such extensions are part of our final conclusion. Let it suffice to say that though location theory has been set forth on the basis of an assumed oligopolistically competitive *f.o.b. mill price system,* the theory applies completely and perfectly to a competitive oligopoly which discriminates in price. Indeed, price discrimination will be shown to be as natural to economic space as is the single price of the competitive spaceless economy. Spatial price discrimination simply alters the distances and price schedules, hence the number of firms. *It does not alter the underlying principles and forces governing the location of a firm.* Moreover, it will be

15. Along this line see Greenhut [3, chap. 6 and app.].

apparent that the welfare aspects of the spatial firm, predicated as they have been in the literature along the lines of a competitive f.o.b. mill pricing oligopolist, apply completely and perfectly to the competitive spatial oligopolist who discriminates in price. But it is clearly time to demonstrate and prove the statements and claims presented above. This will be done initially in Chapter 2 by comparing spatial and spaceless demand and then formulating and comparing simple (spaceless) monopoly price with its spatial counterpart, the f.o.b. mill price.

References

[1] W. Christaller. *Die zentralen Orte in Süddeutschland.* Jena, 1933.
[2] M. F. Dacey. "A Family of Density Functions of Lösch's Measurements on Town Distribution." *Professional Geography,* 1964, 16:5–7.
[3] M. L. Greenhut. *Plant Location in Theory and in Practice.* Chapel Hill, N.C., 1956.
[4] M. L. Greenhut. "Space and Economic Value." *Papers and Proceedings, Regional Science Association,* 1959.
[5] M. L. Greenhut. *A Theory of the Firm in Economic Space,* 2d printing. Austin, 1974.
[6] J. M. Henderson and R. E. Quandt. *Microeconomic Theory.* New York, 1958.
[7] A. Lösch. *Die räumliche Ordnung der Wirtschaft.* Jena, 1944.
[8] E. Mills and M. R. Lav. "A Model of Market Areas with Free Entry." *Journal of Political Economy,* 1964, 72:278–88.
[9] G. Olsson. "Central Place Systems, Spatial Interaction, and Stochastic Processes." *Papers and Proceedings of the Regional Science Association,* 1966, 18:13–45.
[10] W. J. Reilly. *The Law of Retail Gravitation.* New York, 1931.
[11] F. H. Sweet. "An Error Parameter for the Reilly Converse Law of Retail Gravitation." *Journal of Regional Science,* 1964, 5:69–72.

Part 2. Nondiscriminatory Pricing: The Monopoly Case

2. Nonspatial and spatial nondiscriminatory monopoly pricing

I. *Introduction*

This chapter sets forth the relationships which underscore the derivation of *nondiscriminatory* monopoly prices in a spaceless and spatial economy. Its inquiry, based in part on [6], includes determination of the differences between the nondiscriminatory profit-maximizing mill price of the simple monopolist and the counterpart nondiscriminatory mill price of the spatial monopolist. Discriminatory prices are not examined in this chapter because they are not in the form of a single spaceless monopoly price nor the single f.o.b. mill monopoly price of the space economy.

II. *Nonspatial nondiscriminatory monopoly pricing*

Assume a homogeneous set of n buyers, each of whom is possessed of the same identical demand. Then, aggregate as well as individual demand may be defined as

$$\begin{aligned} p &= f(Q) \\ &= f(nq), \end{aligned} \tag{2-1}$$

where p stands for price, Q for aggregate demand, and q for individual demand.

Now, if we assume further that the cost of distance which separates buyers from the seller is negligible, or that all n buyers are located at the same point as the seller's location, the classical theory of simple monopoly requires the firm to equate the additional cost dC of producing the last units of output with the corresponding additional revenue $d(pQ)$.[1] However, the last units of output dQ must be equally divided and purchased by n homogeneous buyers, i.e., $dQ = ndq$, since $Q = nq$. It follows that

1. See J. Robinson [9, chap. 3], and also see A. A. Cournot [2, chap. 5] and A. Marshall [7, chap. 14].

$$\frac{dC}{dQ} = \frac{d(pQ)}{dQ} = \frac{d(pq)}{dq}. \tag{2-2}$$

Thus, marginal cost MC as a function of total output Q must be equated with the additional total revenue derived from the sales of the last additional units of output, call it total marginal revenue, TMR; and TMR must be equated, in turn, with the additional individual revenue derived from each buyer, call it individual marginal revenue, MR_i. In short, the firm's marginal cost must be equated with the marginal revenue it obtains in total and from each buyer.

This fundamental relation may be illustrated in Fig. 2-1 below. Specifically, the figure indicates that total equilibrium output Q_0 is determined at the intersection of the MC curve and the TMR (total marginal revenue) curve. It is also apparent that the equilibrium level of MC, i.e. MC_0, determines the equilibrium level of individual marginal revenue ($MR_i = TMR_0$) and, accordingly, the equilibrium sales ($q = q_0$) to each individual buyer. It follows further that sales to each buyer are not determined in general by the intersection of the MC and MR_i curves. If they were determined at such intersection, the firm would produce the quantity n times this individual demand; but this total clearly implies disequilibrium, viz. MC > TMR, when the MC curve is rising to the right as in Fig. 2-1.[2]

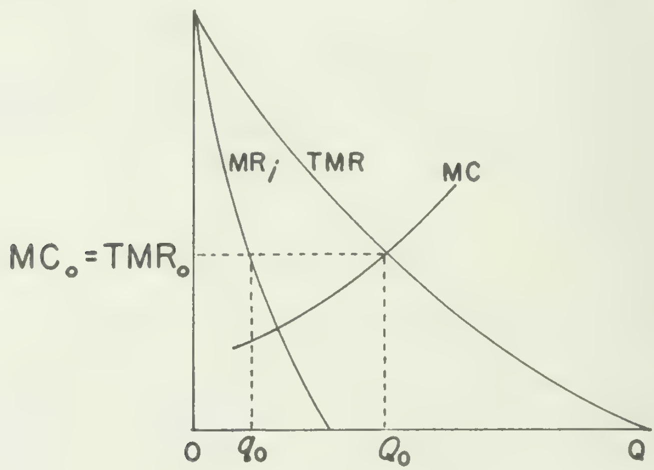


Fig. 2-1 Nonspatial nondiscriminatory monopoly equilibrium.

2. On the other hand, if MC is falling with increasing output, MC < TMR. If MC is horizontal, MC = TMR = MR_i. Only in this last special case may attention be focused on the individual marginal revenue curve in deriving the equilibrium relations in the market. In the other cases, the equilibrium level of MC at its intersection point with TMR must be determined first, not MC with MR_i.

III. *Spatial demand function*

Let us change assumptions by positing an even distribution of buyers along a street such that the first buyer is located at the same site as the seller, the second buyer is located one unit of distance away from the seller, the third buyer is two units of distance away from the seller, and so on. Also assume the freight rate per unit of distance is a positive constant t. Other assumptions being the same, the following set of equations are identifiable:

$$p = f(nq), \tag{2-3}$$

$$p = m + x, \tag{2-4}$$

$$x = tK, \tag{2-5}$$

where q is the quantity demanded at any buying point on the line, p is the delivered price (or the price the buyers are *de facto* required to pay), m the mill price, x the freight rate, and K the distance units from the origin, i.e., from the seller's location to the given buying point.

Equations (2-3)–(2-5) can be reduced by simple substitutions to the general form:

$$q = \frac{1}{n} f^{-1}(m + tK), \tag{2-6}$$

which establishes the quantity of demand in terms of m at any distance K from the seller's site. Equation (2-6) may, therefore, be referred to as the *derived net* individual demand function at a location K distance units from the seller.

Aggregate (market) demand can be obtained, in turn, from (2-6). This derivation requires summation of individual demands as in:

$$Q = \frac{1}{n} \sum_{K=0}^{K_0} f^{-1}(m + tK) \tag{2-7}$$

$$= \frac{1}{nt} \sum_{K=0}^{K_0} f^{-1}(m + tK)t, \quad \forall\, m \ni f(0) - K_0 t \geqslant m > f(0) - (K_0 + 1)t, \text{ and } K_0 = 0, 1, 2, \ldots, n-1,$$

where K_0 is the distance to the most distant buyer. This equation depicts a discrete distribution in which the freight rate per unit of distance t is

high and n small.[3] Of course, if n relates to many buyers evenly distributed along a line, and t is a very small positive constant, e.g., $t = f(0)/n$, then letting $tK = x$, observing the approximation of t by dx, and noting that $K_0 t = f(0) - m$, (2-7) can be rewritten as

$$Q = \frac{1}{nt}\int_0^{K_0 t} f^{-1}(m + x)dx = \frac{1}{nt}\int_0^{f(0)-m} f^{-1}(m + x)dx \qquad (2\text{-}7)'$$

$$= \frac{1}{nt}\int_m^{f(0)} f^{-1}(x)dx = \frac{-1}{nt}\int_{f(0)}^{m} f^{-1}(x)dx, \quad \forall\ m + f(0) > m \geq 0.$$

The first and second derivatives with respect to m are

$$\frac{dQ}{dm} = \frac{-1}{nt} f^{-1}(m) < 0, \qquad (2\text{-}8)$$

$$\frac{d^2Q}{dm^2} = \frac{-1}{nt}\frac{df^{-1}(m)}{dm} > 0, \qquad (2\text{-}9)$$

provided that the assumed gross demand f^{-1} is negatively sloping. Thus, spatial demand is always concave from above, regardless of the concavity or convexity of the individual gross demand curve, which result can also be established for an even distribution of buyers over a plain.[(1)]

IV. *Spatial compared with spaceless demand functions*

The aggregate *spatial* demand function (2-7)′ may be compared with the *spaceless* aggregate demand function (2-1). But first note that the integral $(1/nt)\int_0^{f(0)-m} f^{-1}(m + x)dx$ of (2-7)′ is related to the area of a particular portion of the demand function (2-1). Specifically, it equals the shaded area of the curve given in Fig. 2-2, divided by nt. This constant divisor must be assumed to be equal to or greater than $f(0)$, i.e., $t \geq f(0)/n$, which condition follows, in turn, from the basic assumption that no more than n buyers are part of the monopolist's market.[4]

Curvature of the spatial demand curve differs sharply from that of the

3. The particular value assumed for t signifies that the mill price plus the freight rate applicable to locations greater than K_0 units of distance from the seller exceeds the price intercept value of the demand function. It also signifies that the maximum number of buyers the firm can control over the line market is n.

4. If the freight rate t per unit of distance is assumed to be less than $f(0)/n$, the seller could sell his product to more than n buyers provided that the mill price is very low. But we have assumed n buyers in total. (And see note 3 above. This matter is also discussed further in Chapter 4.)

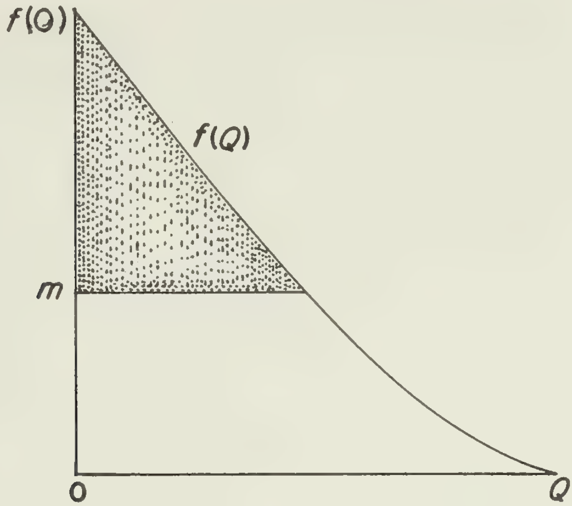


Fig. 2-2 Integrating the inverse of $f(Q)$ from m to $f(0)$.

spaceless demand curve. To understand this important relationship, assume for the moment the linear spaceless aggregate demand curve

$$p = f(0) - aQ, \quad \text{or alternatively} \tag{2-10}$$

$$Q = \frac{(f(0) - p)}{a} \tag{2-10$'$}$$

Then via (2-7)′ the spatial aggregate demand takes the form

$$Q = \frac{1}{nt} \int_{m}^{f(0)} \frac{(f(0) - x)}{a} dx = \frac{(f(0) - m)^2}{2af(0)} \tag{2-11}$$

provided $nt = f(0)$, where m is the spatial firm's mill price as in (2-4). This result conforms to the fact that when buyers are distributed evenly along a line, the average distance in the market can easily be shown to be one-half of the distance from the seller to the most distant buyer.[II] If a constant freight rate per unit of distance is assumed, average freight cost is one-half the maximum freight cost the seller can charge the most distant buyer. This particular result signifies that if the spatial mill price m is zero, the spatial quantity intercept, via the linear spaceless demand function $Q = (f(0) - p)/a$, equals $f(0)/2a$ rather than $f(0)/a$; correspondingly, the average (delivered) price is $f(0)/2$ greater as a result of freight cost than in the spaceless case. And when mill price is $f(0)$, no sale can be made to any distant buyer(s). It follows that spatial demand coincides with spaceless demand at the price intercept point. Spatial

impacts thus become increasingly distinctive as mill price is lowered from $f(0)$ to 0.

Effective demand in the space economy therefore consists of two multiplicative components, $(f(0) - m)/a$ and $(f(0) - m)/2f(0)$, where $0 < (f(0) - m)/2f(0) \leq 1/2$ for m ranging from $f(0)$ to 0. Observe with respect to the factor $(f(0) - m)/2f(0)$ that the spatial buyer's effective demand is sharply circumscribed as a result of the freight cost to his site. That is to say, as mill price is raised above zero, not only is the total quantity demanded decreased according to the slope a of the linear demand function, but the most distant buyer who previously was part of the seller's market (i.e., the buyer at the economic distance $f(0)$) finds the delivered price excessive; that is, $m > 0 +$ freight cost tK_0 ($= f(0)$). Hence he drops out of the market. The $(f(0) - m)/2f(0)$ quantity demanded in the market at $m > 0$ stands, accordingly, in the proportion $(f(0) - m)/2f(0)$ of the greatest possible quantity that would be purchased, i.e., the spaceless demand quantity $(f(0) - m)/a$. Since $(f(0) - m)/2f(0)$ does not change demand linearly with m, *and* $(f(0) - m)/2f(0) < \frac{1}{2}$ when $m > 0$, concave spatial demand appears as in Fig. 2-3. (Note that Q^* is less than $Q/2$.)

To sum up, when mill price $m = f(0)$, spatial and spaceless quantities are both zero. In other words, spatial demand $Q = (f(0) - m)^2/2af(0)$ is 0, and spaceless demand $Q = (f(0) - m)/a$ is also 0, where in the latter expression m is substituted for p to effect easy comparisons. When $m = 0$, however, distance has its maximum impact on spatial demand, as the firm can sell to the potentially most distant market point (i.e., the most distant effective demander in the market). The spatial quantity demanded when $m = 0$ is given by $Q = (f(0) - m)^2/2af(0) = f(0)/2a$, which

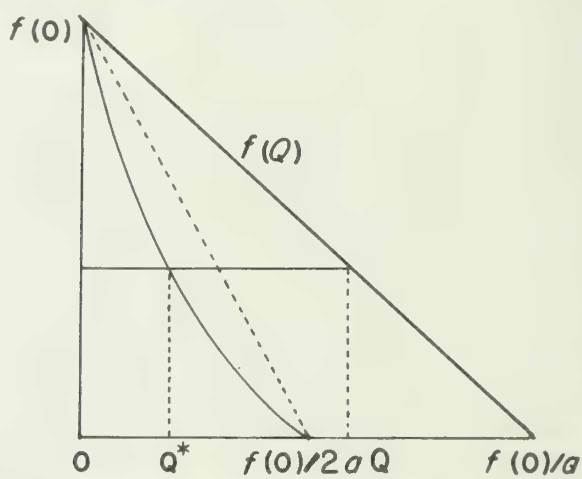


Fig. 2-3 Deriving spatial demand Q^* from spaceless demand Q.

is one-half the spaceless case of Q equal to $f(0)/a$.[III] It is the factor $(f(0) - m)/2f(0)$ which guarantees concavity of the spatial demand curve, as d^2Q/dm^2 is positive in the present model, being equal to $1/af(0)$.

The relation scored above between a linear spaceless demand curve and the corresponding spatial demand is illustrated in Fig. 2-3. The concave curve in the figure depicts the spatial aggregate demand under the assumption that buyers are distributed along a line market; the spaceless aggregate demand is described by the linear curve.

When n buyers are distributed evenly over a plain, aggregate demand is

$$Q = \frac{1}{a\pi f(0)^2} \int_0^{2\pi} \int_0^{x_0} (f(0) - m - x)x dx\, d\theta, \tag{2-12}$$

where x_0 is the freight cost limit and $\pi f(0)^2$ is the area of a circle. Thus $\pi f(0)^2$ is counterpart to $f(0)$, i.e., the area of a line market. By elementary manipulations, we obtain

$$\begin{aligned} Q &= \frac{1}{a\pi f(0)^2} \int_0^{2\pi} \int_0^{f(0)-m} [(f(0) - m)x - x^2] dx\, d\theta \qquad (2\text{-}13) \\ &= \frac{2\pi}{a\pi f(0)^2} \left[\frac{(f(0) - m)^3}{6}\right] \\ &= \frac{(f(0) - m)^3}{3af(0)^2}. \end{aligned}$$

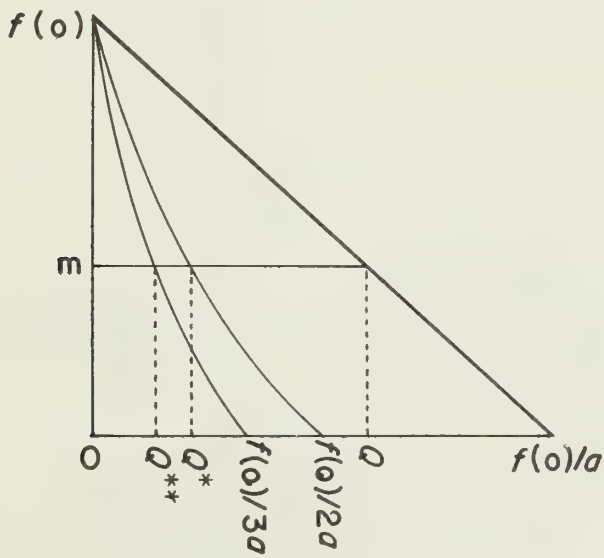


Fig. 2-4 Spatial demands under alternative buyer distributions.

When mill price is $f(0)$, $Q = (f(0) - m)^3/3af(0)^2 = 0$. When mill price is 0, it follows that $Q = f(0)/3a$. The quantity demanded is just one-third of spaceless demand and two-thirds that of the line case, as shown in Fig. 2-4. Also it is clear again that the concavity of the spatial demand is guaranteed.

V. *Spatial vs. nonspatial nondiscriminatory pricing*

It has been shown how market demand in the space economy is derivable from nonspatial market demand. Significantly, the former is smaller in magnitude than the latter. But this fact does not necessarily imply that the equilibrium spatial f.o.b. price is less than the nonspatial price. To appreciate the differences involved, note that the spatial monopolist equates his marginal cost of production with the *spatial* marginal revenue derived from the spatial market demand function. It follows, for example, that if the spaceless demand curve is linear and the marginal cost curve constant, the spatial price will be lower than the spaceless price in equilibrium. To see this result, let the spaceless aggregate demand be specified again as $p = f(0) - aQ$ or $Q = (f(0) - p)/a$. Then the spatial aggregate demand is given—as before—by $m = f(0) - \sqrt{(2af(0)Q)}$ or $Q = (f(0) - m)^2/2af(0)$, when buyers are distributed along a line, and by $m = f(0) - (3af(0)^2Q)^{1/3}$ or $Q = (f(0) - m)^3/3af(0)^2$ in the event of buyer distribution over a plain.

The spaceless and spatial marginal revenue functions MR_i (i = a, b, c) thus are respectively specifiable as

$$\mathrm{MR_a} = f(0) - 2aQ, \tag{2-14}$$

$$\mathrm{MR_b} = f(0) - \frac{3}{2}(2af(0)Q)^{1/2}, \tag{2-15}$$

$$\mathrm{MR_c} = f(0) - \frac{4}{3}(3af(0)^2Q)^{1/3}, \tag{2-16}$$

where subscripts a, b, and c refer to buyer distributions at a point, along a line, and over a plane respectively. Let MC = c and equate it with the spaceless and spatial marginal revenues to obtain[IV]

$$c = f(0) - 2aQ \quad \rightarrow \quad Q = \frac{(f(0) - c)}{2a}, \tag{2-17}$$

$$c = f(0) - \frac{3}{2}(2af(0)Q)^{1/2} \quad \rightarrow \quad Q = \frac{(2/3)^2(f(0) - c)^2}{2af(0)}, \tag{2-18}$$

$$c = f(0) - \frac{4}{3}(3af(0)^2Q)^{1/3} \quad \rightarrow \quad Q = \frac{(3/4)^3(f(0)-c)^3}{3af(0)}. \tag{2-19}$$

Substituting these equilibrium values respectively into equations (2-10), (2-11), and (2-13) provides the profit-maximizing equilibrium spaceless and spatial prices[V]

$$p \equiv m_a = \frac{(f(0)+c)}{2}, \tag{2-20}$$

$$m_b = \frac{(f(0)+2c)}{3}, \tag{2-21}$$

$$m_c = \frac{(f(0)+3c)}{4}. \tag{2-22}$$

It follows that $p > m_b > m_c$ as long as c is so constrained that $f(0) > c \geqslant 0$, which is, of course, an economically relevant constraint.[5] Equilibrium mill price under monopoly thus depends upon how buyers are distributed over economic space. In particular, a relatively greater concentration of buyers at nearby points generates higher mill prices than would the opposite distribution.

It is possible when MC is declining rapidly for the equilibrium spaceless price to be lower than the equilibrium spatial price. This special result stems from the high-level intersection of a falling MC curve with the spatial aggregate marginal revenue curve; in turn, its intersection in this event with the spaceless aggregate marginal revenue curve occurs at a significantly low level.

The fall (or rise) in equilibrium spatial price below (or above) the spaceless price is illustrated in Fig. (2-5) by the values p, m_0 and m_1. The heavy lines in Fig. (2-5) stand for the spaceless aggregate demand curve and its corresponding marginal revenue curve; the two light curves, also originating at $f(0)$, apply respectively to some spatial aggregate demand curve (actually in the figure the curve applicable to the line distribution) and its associated marginal revenue curve. Observe that the alternative marginal cost curves, MC_0 and MC_1, cut the spatial marginal revenue

5. The fundamental relation given above was demonstrated years ago by M. L. Greenhut [5]. In similar key (and for use later in the book) we may consider the equilibrium mill prices under the alternative spatial distributions of buyers as a function of K_0^* where K_0^* is the boundary limit of the firm's sales radius, and the asterisk is added here and henceforth to stress the assumption that sales are extended to the profit-maximizing distance in the seller's market. Specifically, using b in place of $f(0)$ and assuming a unitary freight rate, we shall obtain in the Appendix to this chapter: (i) $m_b = (b+c)/2 - K_0^*/4 = c + K_0^*/4$, and (ii) $m_c = (b-c)/2 - K_0^*/3 = c + K_0^*/3$. Direct comparison of (i) and (ii) also points to the relation $m_b < m_c$ established in the text. And cf. [8] and [1].

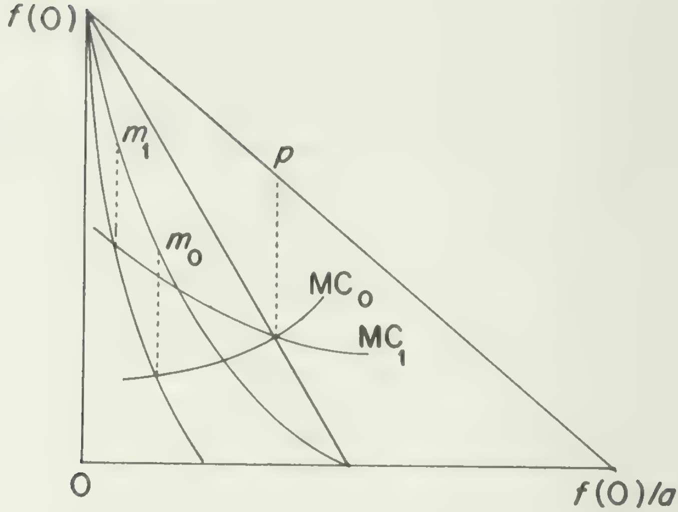


Fig. 2-5 Spatial and spaceless prices under alternative cost conditions.

curve at different points. These intersections yield different equilibrium price levels, namely, m_0 and m_1.

It is clear that the price effects of space depend, in general, on cost conditions, viz., the shape of the MC curve. If the MC curve is horizontal or rising, the spatial price will be lower than the spaceless price. If, on the other hand, substantial economies of scale prevail and the MC curve falls rapidly, the effect of economic space is to raise prices. Minimization (or economization) of the *cost* of distance is clearly required to enjoy the "fruits" of scale economies, i.e., lower mill price through increased output. In the more likely case of a rising or horizontal MC curve, a reduction in the cost of distance helps increase total output but raises *mill* price too, as the spatial demand curve and hence the mill price come closer to the demand curve and the price applicable to the spaceless economy. Delivered prices to buyers at substantial distances from the seller would, of course, be less than they were before the cost of distance had fallen.

Though simple spatial monopoly pricing is obviously more complicated than spaceless monopoly price determination, a close relation underscores the two, since the spatial f.o.b. mill price is directly counterpart to the simple spaceless monopoly price. Chapters 3 through 6 will indicate, however, that the subject of monopolistic price discrimination over economic space is much more complicated than spaceless monopoly price discrimination. Indeed, many "remarkable" implications of spatial discriminatory pricing will be uncovered which contrast sharply with

results derived in studies of spaceless discriminatory price. These differences alone are enough to indicate the need for studies of spatial prices.

Appendix: *Alternative approach to mill price*

We assumed in the text that a given number of buyers are dispersed along either a line market or over a plain. Correspondingly, the net demand density at a point on the line market was defined as $(f(0) - m - tK)/af(0)$ whereas, in the case of an even distribution of buyers over a plain, the demand density at a point was given by $(f(0) - m - tK)/a\pi f(0)^2$. Demand density at a point was, in other words, assumed to vary as the distribution of buyers was changed.

This appendix will demonstrate that equilibrium mill prices are invariant to different demand densities at a point. In particular it assumes that a given net demand density of the form $(b - m - K)/a$ applies regardless of buyer distribution, where b is used interchangeably with $f(0)$ and the freight rate t is assumed to be unity. Aggregate demand for the linear market is then given by

$$Q = \int_0^{K_0} (b - m - K)\left(\frac{1}{a}\right) dK \tag{2-23}$$

$$= \left[(b - m)K_0 - \left(\frac{K_0^2}{2}\right)\right]\left(\frac{1}{a}\right),$$

where K_0 stands for the boundary limit point to the firm's sales radius.

If marginal cost of production c is constant, profits of the firm are specifiable as

$$\Pi = (m - c)Q - F \tag{2-24}$$

$$= \frac{(m - c)[(b - m)K_0 - (K_0^2/2)]}{a} - F,$$

where F stands for fixed cost of production. The first-order condition for profit maximization is

$$\frac{d\Pi}{dm} = \frac{[b + c - 2m - (1/2)K_0]K_0}{a} = 0, \tag{2-25}$$

$$m = \frac{(b + c)}{2} - \frac{K_0}{4}. \tag{2-26}$$

Since the natural limit to the seller's market is defined by $b = m + K_0$, equation (2-26) above reappears as

$$m = \frac{c + K_0}{2}, \tag{2-26)'}$$

$$m = \frac{(b + 2c)}{3}, \tag{2-26)''}$$

as derived in (2-21) above.

In the case of a continuous distribution of buyers over a plain, the aggregate demand is given by

$$Q = \int_0^{K_0} 2\pi K\,(b - m - K)\left(\frac{1}{a}\right)dK \tag{2-27)}$$

$$= \frac{\pi\{(b - m)\,K_0^2 - (2/3)\,K_0^3\}}{a}.$$

Profits then are specifiable as

$$\Pi = (m - c)Q - F \tag{2-28)}$$

$$= \frac{\pi(m - c)[(b - m)\,K_0^2 - (2/3)\,K_0^3]}{a}.$$

The first-order condition for maximization is

$$\frac{d\Pi}{dm} = \frac{\pi\{(b + c - 2m)K_0^2 - (2/3)K_0^3\}}{a} = 0, \tag{2-29)}$$

$$m = \frac{(b + c)}{2} - \frac{K_0}{3}. \tag{2-30)}$$

Again, via $b = m + K_0$ this solution may be rewritten as

$$m = \frac{c + K_0}{3}, \tag{2-30)'}$$

$$m = \frac{(b + 3c)}{4}, \tag{2-30)''}$$

as derived in (2-22) above.

The foregoing analysis does not imply that population density is ir-

relevant in determining the level of equilibrium mill price. Manifestly, if marginal cost is rising (falling) to the right, then mill price will be higher (lower) when buyer distribution involves greater population density.

References

[1] M. Beckmann. *Location Theory.* New York, 1968.
[2] A. A. Cournot. *Researches Into the Mathematical Principles of the Theory of Wealth.* Translated by N. T. Bacon. 1897.
[3] K. G. Denike and J. B. Parr. "Production in Space, Spatial Competition, and Restricted Entry." *Journal of Regional Science,* 1970, 10:49–63.
[4] C. A. Gannon. "Fundamental Properties of Löschian Spatial Demand." *Environment and Planning,* 1971, 3:283–306.
[5] M. L. Greenhut. *Plant Location in Theory and in Practice.* Chapel Hill, N.C., 1956.
[6] M. L. Greenhut, M. J. Hwang, and H. Ohta. "Spatial Aspects of Economic Development." *Journal of Regional Studies* (The Japanese Association of Regional Science), 1973, 3:239–54.
[7] A. Marshall. *Principles of Economics.* 8th ed. 1961.
[8] E. Mills and M. Lav. "A Model of Market Areas with Free Entry." *Journal of Political Economy,* 1964, 72:278–88.
[9] J. Robinson. *The Economics of Imperfect Competition.* 1933.

3. A review of theories of nonspatial and spatial monopoly price discrimination

I. *Introduction*

From this point forward we shall employ models which go beyond comparing the mill price levels of spatial and spaceless monopolists, as answer is sought to the question whether the underlying price policy of the firm in economic space is actually similar to or quite different from that of the classical nonspatial firm. The models to be employed point to a compelling need for including the concept of economic space (i.e., the distance variable) in basic economic theory. Indeed, the mere fact that a proof is ultimately given that competitive firms in economic space often *will discriminate in price* suggests, by itself, the extent to which spatial conceptions and related theory deviate from the classical system of economic thought.

There are additional points of difference between the classical and spatial analyses of the price policies of firms. To mention another at this time, recall from classical economic theory that in the absence of legal restraints, arbitrage resale conditions, and the like, that it is essentially only the shape of the demand curve which was said to be important in determining whether a monopolist discriminates against certain buyers in his pricing policy. Under the same *ceteris paribus* pound, spatial economics concludes that the magnitudes of freight costs in addition to production costs play decisive roles along with the demand curve shape in determining the existence (nonexistence) of price discrimination.

The differences mentioned above are but two of many contrasts which underscore the theory of pricing of nonspatial and spatial *firms*. For the remainder of Part II of this book, however, attention is confined to the pricing patterns of the spaceless and spatial *monopolist;* in turn, the price patterns of *competitive firms* are left to Parts III and IV of the book.

Consider two other points of contrast: first, the relations between outputs and pricing patterns. In particular, which pricing system yields more outputs? The classical economic (nonspatial) answer to this question was that the comparative sizes of outputs depend on the shape of the

demand functions. It will, however, be shown (Chapter 4) that discriminatory pricing yields more outputs under spatial monopoly than f.o.b. mill pricing *regardless of the shape of the demand functions.* Moreover, it will be shown (Chapter 8) that spatial competition involves in general the basically same relations, again *regardless of the shape of the demand functions.* As with the pure matter of price patterns, the question of outputs of *competitive firms* in spaceless and spatial economic theory is reserved to Parts III and IV of the book.[1]

As the second point, a relation exists in the space economy, between the sizes and shapes of market areas and the applicable price and market structures. Part IV of the book thus demonstrates this added element by determining the functional relationships between the *price and output* practices of spatial firms and the sizes and shapes of their market areas. In fact, the exact mapping (or function) which ties the one to the other is specified in the final chapter of the book!

II. *A review of classical theories of price discrimination*

J. Dupuit

An early penetrating analysis of alternative pricing techniques can be found in Dupuit [7], where he discussed discriminatory pricing vis-à-vis simple monopoly pricing.[2] He claimed that if monopoly power of some degree exists and if "buyer estimates" by the seller are different, the firm will sell a given product for different prices.[3] Such a practice is necessarily more profitable than the practice of quoting only one price for all buyers. That discriminatory pricing also yields more outputs and therefore more welfare or less "lost utility" is the central conclusion drawn by Dupuit.

Dupuit is one of the promoters of diagrammatic presentations in eco-

1. It would be desirable to test empirically the several conclusions to be derived in this book. However, our emphasis is chiefly on creating theory. Empirical verification of the theory to be set forth herein must be reserved for other writings.

2. Professor R. B. Ekelund Jr. [12, p. 269] observes that Dupuit is rather well known in connection with utility analysis and, in particular, the concept of consumer surplus, but he is not connected generally with the theory of price discrimination. Cf. Marshall [25, pp. 85, 394]; Hicks [17, pp. 38–39]; Schumpeter [31, pp. 838–39]. Even Mrs. Robinson [29], who reopened the discussion of price discrimination almost a century later, did not mention Dupuit. In contrast, Walras [40, p. 443] and Edgeworth [9, p. 441] recognized Dupuit's contribution to the subject.

3. Cf. Dupuit [7, pp. 88–89], where he states: "The variable, yea mobile nature of the value of utility is indeed well known to business men. . . . That is what lies behind all transactions which are sheltered from competition." "Why are there two different prices for the same service? Because the poor man does not attach the same value to crossing the bridge as the rich man does."

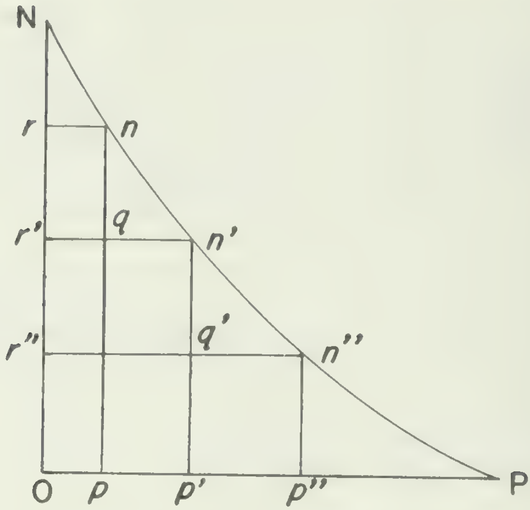


Fig. 3-1 Lost and remaining utilities

nomic analysis.[4] Fig. 3-1 follows in his tradition by measuring price on the horizontal axis and quantity demanded, or the number of buyers, on the vertical scale [7, p. 108]. If, for simplicity, it is assumed further with Dupuit that marginal cost is zero,[5] the profit-maximizing output and price of the monopolist can be obtained from the equation $d(pq)/dp = 0$. Accordingly, let the rectangular area $Op'n'r'$ in Fig. 3-1 represent the greatest possible profits of the firm. Then the area $p'Pn'$ provides what Dupuit calls the utility remaining to the consumers.[6] The lost utility is $r'n'N$.

According to Dupuit, the "lost utility" which takes place in a market is influenced by the degree of monopoly power possessed by a firm. In the limiting case of zero costs and no monopoly power, lost utility would vanish completely. What Dupuit actually is saying is that lost utility is attributable to the negative slope of the demand curve, or, more fundamentally, the existence of finite elasticit*ies* of demand. But note, his concept of monopoly power does not necessarily imply monopoly profits;

4. Schumpeter [31, pp. 383–89] states that "such merit as there was in the rediscovery of the marginal utility principle is Jevons'; the system of general equilibrium . . . is Walras'; . . . the consumers' rent is Dupuit's; the diagrammatic method of presentation is also Dupuit's." The last statement of Schumpeter's, however, is not valid, for Cournot [4] had elegant presentations of diagrammatic analyses before Dupuit.

5. We interpret this to be an implicit assumption in Dupuit's theory, as he only stipulated the existence of fixed cost [7, p. 107]. Without our interpretation of Dupuit's assumptions, his *revenue* maximization formula [7, p. 108] would not imply *profit* maximization.

6. Hicks [17, pp. 38–39] stated that this concept is equivalent to Marshall's "consumers' surplus" (Marshall [25, p. 103]) under the assumption that the marginal utility of money is constant. However, it is shown in the Appendix to this chapter that such is not the case.

monopoly power, and hence lost utility, could exist even if there were no monopoly profits available at all.[7] It follows that if "lost utility" happened to decrease as monopoly profits increased, Dupuit could say that monopoly power decreased.[8] Correspondingly, substitution of price discrimination for simple monopoly pricing could reflect a decrease in lost utility and at the same time an increase in monopoly profits as monopoly power decreases. This correspondence and the related increase in output would be viewed as follows.

Suppose there are three sets of consumers, each separable from the others so that different "buyer estimates" apply. In particular, assume in Fig. 3-1 that one set of buyers possesses the demand formed by $OPn''r''$, the next set the demand polygon $r''n''n'r'$, and the next (and weakest) buyers the demand polygon $r'n'nr$. Alternatively, we might assume a finitely limited, perfectly inelastic individual demand such that the set of buyers with the strongest demands (i.e., the Or'' set of buyers) could afford at least as high a price as $r''n''$, the next buyer set $r''r'$ would buy at the high price $r'n'$, and the $r'r$ set of buyers (i.e., the buyers with weakest demands) would purchase at price rn. Let resale possibilities be ruled out. Then it is possible to increase outputs and hence eliminate lost utility by using three different price (output) combinations, namely, prices $r''n''$, $r'n'$, and rn. Monopoly profits or the tax yield in the case of perfectly inelastic individual demands would, according to Dupuit [7, p. 108], be the sum of the three rectangles $Op''n''r'' + r''q'n'r' + r'qnr$. Given the three demand polygons under the *initial assumption in this paragraph* (for example, the demand of the first group is formed by the polygon $OPn''r''$), the utility remaining to the consumers would be the sum of the three (quasi) triangles $p''Pn'' + q'n''n' + qn'n$;[9] hence, the lost utility would amount to rnN, i.e., the total utility OPN less the sum of the utility received by the monopolist and the consumer. If a full set of

7. For example, a Chamberlinean tangency is said to involve monopoly power but no monopoly profits. See Chamberlin [2, p. 84]. In this connection, Lerner's degree of monopoly also has nothing to do with monopoly profit [21, p. 169].

8. This relationship may be paradoxical. The paradox stems directly from the semantics of present day theorists. Thus on the one hand monopoly power is related to elasticity of demand and on the other hand to monopoly profit. Obviously, the two criteria are completely independent.

9. This proposition holds because Dupuit assumed perfectly inelastic individual demand curves. In particular, he conceived [7, pp. 108–9, 8, pp. 7–9] of the following situation: If the toll on a bridge is very high, no one crosses the bridge. If the toll is reduced to some extent, the richest members of the village will cross the bridge, but not others. If the toll is lowered still further, the next set of villagers will cross the bridge rather than stay on the one side or reach the other side by swimming across the stream. Any lowering of price thus elicits an increase in the number of customers using the facility. Dupuit conceived, therefore, of a separation of customers into, say, three groups, "the rich, the moderately well-off, and the poor" [7, p. 88], with their respective aggregate demands given as in the text above. Most important, this same proposition is approximated if we assume a constant *and* independent marginal utility of money so that a Marshallian demand curve subject to this constraint is drawn. See the Appendix to this chapter for details.

buyers ON exists, and each and every buyer is completely separable, monopoly profits would be the triangle OPN and the lost utility as well as the utility remaining to the consumers (i.e. consumers' surplus) would both be zero. This is exactly the case of what is called "discrimination of the first degree" or "perfect discrimination" by Pigou [28, p. 277] and Robinson [29, p. 187], respectively.[10]

The analysis above suggests that a monopolist will try to discriminate among his buyers just so long as it is technically possible to do so. However, Dupuit implied that a public monopoly might follow a policy of constrained discrimination. Certain basic problems then arise, in particular the problem whether the largest possible profit is earned and whether the loss of utility is reduced to a minimum; see Dupuit [8, p. 31].

According to Ekelund [12, p. 276], Dupuit apparently believed that governments should strive to maximize consumers' surplus. However, Ekelund also noted [12, p. 273] that Dupuit believed that "discrimination was desirable only if it increased quantity over that obtained under a single price simple monopoly system, for only in that event would *utilité perdue* be reduced." Manifestly, the two criteria, i.e., maximization of consumers' surplus and maximization of outputs, are not necessarily equivalent, for the reason that maximizing output by complete price discrimination implies minimizing *not maximizing* consumers' surplus. Of course, the minimized consumers' surplus can be altered (or offset) by redistributing the maximized yield.

Dupuit's contribution to the theory of price discrimination is important from the standpoint of public policy as well as pure theory, although his writings on this subject were for a long while almost completely overlooked. They were, in fact, generally ignored up to F. Y. Edgeworth, who claimed in 1910 [9, p. 441] that Dupuit was the earliest and "highest authority on the theory of price discrimination."[11] Incidentally, Edgeworth's own contribution to price discrimination stems from his analysis of the interdependence of demand under conditions of price discrimination.[12]

A. C. Pigou

In 1920 Pigou presented another view of price discrimination by classifying it according to three types – first, second, and third degrees

10. Ekelund [12, p. 271] interprets Dupuit's model as involving another kind of discrimination, which he designates as the third degree of discrimination. In effect, we shall see that there is no conceptual difference between the two interpretations so far as Dupuit's analysis (and theory) is concerned because only one unit of service is supposed to be bought by each buyer.

11. However, it must be pointed out that Walras [40, p. 443], before Edgeworth, did observe (or more precisely mentioned) Dupuit's contribution.

12. See [9, pp. 441–43]. In a similar vein, Simkin [32, pp. 4–5] discussed interdependence in his general equilibrium theory of discrimination. Also see Stigler [37, p. 211].

of price discrimination [28, p. 279]. We describe these alternative types of price discrimination in some detail in the Appendix to this chapter. For the moment, let it suffice to mention these points: (1) The first degree of discrimination involves different prices for each unit of commodity sold, as in Dupuit's example of buyers possessed of differently limited perfectly inelastic demands. In this case of first-degree discrimination, no consumer surplus is left at all as the monopolist appropriates everything for himself. (2) The second degree is (and will be shown to be) a less complete form of the first degree. Ostensibly it involves discrimination by groups with some consumer surplus remaining such as areas $p''Pn''$, $q'n''n'$, and $qn'n$. (3) Third-degree discrimination is significantly different. It involves a separation of markets, each of which has its own but different demand function. As will be stressed in Chapter 4, profits are maximized when marginal revenues are equated with the relevant marginal costs *in each market.* This maximization involves charging different prices in the two markets, with the price in the market of lowest demand elasticity (as defined at the simple monopoly price level) being the highest under discrimination. In general, Pigou stressed the case of third-degree discrimination under the assumption of linear demand functions.

Joan Robinson

Besides synthesizing previous writings on price discrimination, Joan Robinson [29, p. 179–201] established the conditions under which price discrimination yields more (or less) outputs as a whole than does simple monopoly pricing.[13] She observed that total output would be greater or less under monopoly price discrimination depending upon whether the more elastic of the demand curves in the separate markets is more or less concave than the less elastic demand curve; in turn, outputs will be the same if the demand curves are straight lines or if the concavities are equal [29, p. 190].

To appreciate Mrs. Robinson's conclusions, it is obvious, for example, that price discrimination in the presence of parallel linear demands in each market would increase the firm's production for the weaker market (i.e., the market whose demand curve is leftward and downward of that of the other market) while decreasing its output in the stronger market *by the same amount.* But if the weaker market has a concave rather than linear demand, the MR, MC equality would cause output to be greater under concave demand in the weaker market than what it would be under linear demand. These relations may readily be understood with the aid

13. Before Mrs. Robinson's writings, a controversy had taken place between Viner [39] and Yntema [41] as to whether or not discrimination could ever benefit members of the stronger market. As Mrs. Robinson pointed out, this controversy would not have arisen if the writers had been familiar with Pigou's analysis.

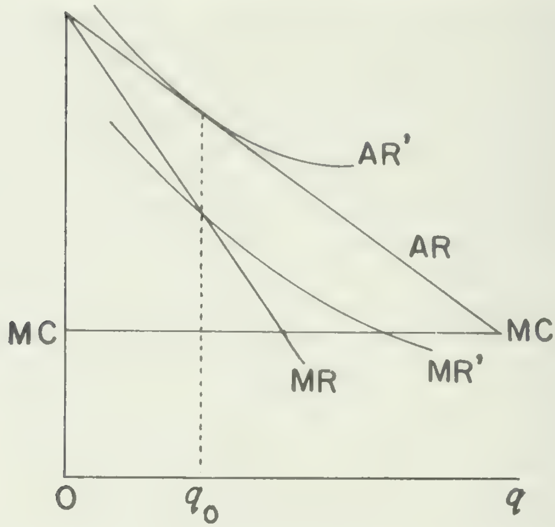


Fig. 3-2 Demand concavity and output effect of price discrimination.

of Fig. 3-2, where the weaker market is characterized alternatively by the concave and linear demands given in the figure. To view the results, assume that q_0 is the output which would be shipped to the weaker market under conditions of *nondiscriminatory* simple monopoly pricing. Under price discrimination, firms equate MR with MC in each market. The MR′ and MR curves thus apply to the weaker market, given the concave and linear demand curve alternatives assumed above. Their intersection with marginal cost produces a lower price and greater output in that market. But the increase in output for the weaker market is greater under price discrimination when its demand curve is concave and not linear, since elasticity decreases less rapidly on this curve as price decreases compared with the decrease in elasticity at lower prices along the linear curve. Total output for all markets is therefore always greater when the demand curve in the weaker market is concave relative to the demand curve in the stronger market, provided further that the weaker market is also characterizable as having the more elastic demand at the simple monopoly price. The converse holds when price is raised in the weaker market, because elasticity is less in that market. Mrs. Robinson also noted that opposite arguments hold for the convex case.[14] In effect, she was contending that demand conditions determine the direction of output under price discrimination.

It can, however, be shown that cost conditions also affect the direction

14. For a more detailed discussion of Robinsonian analysis of outputs under discrimination, see Robinson [29, pp. 190–93] and Edwards [11, pp. 166–71].

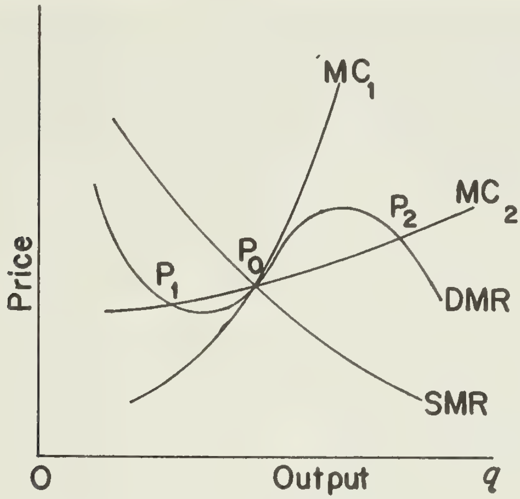


Fig. 3-3 The shapes of SMR, DMR, and MC and output effect of price discrimination.

of output under price discrimination. To understand our thesis, let us refer to the marginal revenue curve of a discriminatory monopoly as DMR and that of a nondiscriminatory (simple) monopoly as SMR. Consider, in turn, the SMR and DMR curves of Fig. 3-3, originally used by Mrs. Robinson [29, Fig. 65, p. 198].[15] Outputs produced by a nondiscriminating or a discriminating monopolist will be the same if their marginal cost curves are MC_1, *but different if their curves are* MC_2. The point of stable equilibrium is either P_1 or P_2, and not P_0 under discriminatory monopoly when the marginal cost curve is MC_2. P_0 is unstable because MC_2 cuts DMR from above as one reads from left to right. Now, all of this is not to say that Mrs. Robinson did not recognize the relevance of the *position* of the MC curve in determining the direction of outputs under alternative pricing techniques. In fact, she observed that if the MC curve intersects the DMR and SMR curves "where they cut each other," the output under discriminatory monopoly is the same as that under simple monopoly; see [29, p. 196]. Output would, therefore, be influenced by MC curves which intersect DMR and SMR curves at different points. But we see in Fig. 3-3 that even if the MC curve (i.e., MC_2) intersects DMR and SMR "where they cut each other," the outputs are different. It is her related statement [29, p. 195] — "If the effect of discrimination is to leave the total output unchanged, it will make no difference whether costs are rising, falling, or constant" — which either misleads or is insufficient.

15. The shapes of the *spatial* DMR and SMR curves will be derived in Chapter 4.

Cost conditions help determine whether or not output is increased (or decreased) under price discrimination compared to simple monopoly pricing.[16] Thus all one may say about outputs according to nonspatial price-discrimination theory is that typically discriminatory prices yield more output, a statement which assumes that even if convex demand exists, the necessary offsetting relationships leading to greater output may well prevail. In fact, Mrs. Robinson asserted that such relationship is "likely to be common" [29, p. 201]. To repeat (and again paraphrase Mrs. Robinson), price discrimination in the spaceless economy generally increases output.

Notwithstanding the penetrating analyses and specific findings of Dupuit, Pigou, and Mrs. Robinson, besides those of Edgeworth [9], Samuelson [30], Stigler [37], Simkin [32], and Edwards [11], who – for reasons of objectives – need not be discussed individually herein, the classical (nonspatial) literature on price discrimination has been obtuse. But the literature on spatial price discrimination has been correspondingly obtuse, even though few alternative possibilities appear to mark the generalizations that have thus far been set forth in this body of economic thought.

III. *A review of theories of spatial price discrimination*

One may wonder why the term concavity (or convexity) of demand has been used rather than elasticity. Surely the form (or shape) of a demand curve can be expressed in terms of the elasticity of demand. However, though the form of the demand function helps determine the elasticity of demand or, more precisely, the rate of change of elasticity, one may not go directly from elasticity into form. For example, mere knowledge that elasticity is unity does not carry forward in providing a unique demand function, because unitary elasticity may occur at different prices on innumerably many demand functions. It is for this reason that the shape (concavity) of the demand curve is a more generally applicable term in establishing certain principles and facets (e.g., output effects) of price discrimination. This is why Mrs. Robinson and P. Samuelson, among others (see [29], [30], [11]), have used this particular reference. Indeed, in Dupuit's model, there is no advantage in comparing elasticities when each buyer is assumed to possess a perfectly inelastic demand curve, i.e., $e = 0$, and the monopolist nevertheless is shown to gain from

16. In this connection, Samuelson [30, p. 45] also stated without proof that the curvatures of demand curves would determine whether output is greater or smaller under discrimination. However, his analysis implies that the cost curves are also generally relevant in determining the direction of a change in output.

discrimination in price. So the term elasticity sometimes is inadequate in discussions of the prices resulting under price discrimination, though generally such reference base is sufficient. In fact, it will be seen that the concept of elasticity actually underscores the literature of spatial price discrimination theory perhaps more than the classical references to shapes.

A general price discrimination constraint in economic space

Price discrimination is possible only if the resale of a good by a "favored" buyer to an "unfavored" buyer is impractical, impossible, or otherwise unlikely to occur. This necessary requirement for price discrimination will tend to prevail in a space economy whenever the discrimination is applied against the buyers located more proximate to the seller than the other buyers. By introducing the spatial dimension, i.e., the distance variable, an interesting perspective is added to the subject of price discrimination. But though this condition might suggest that spatial price discrimination should have become an integral part of the subject at a relatively early date in the development of price discrimination theory, such was not the case. The analysis of spatial price discrimination followed that of nonspatial price discrimination by about fifty years; and even today it is largely discussed by only a rather small group of economists.

Analysis of spatial price discrimination was first set forth by Hoover [18], then by Singer [33], Leontief [20], Smithies [35], Lösch [23], Dewey [5], and Greenhut [15]. In time, the subject gained a "richness" reflective of the increased "realism" and empirical reference points which are intrinsic to models relating to economic space. When one considers the fact that basing-point and related systems are in fact discriminatory against buyers at selected locations, the importance of the subject is clear. To say nothing else, the place of this pricing system in the economic development of the United States has been well documented in the writings of economic historians.

E. M. Hoover

Although E. M. Hoover's theory will be set forth again in Chapter 5, his contribution to the theory of spatial price discrimination must be reviewed here not only because he was the first authority to apply the new variable, i.e., economic space, to the subject, but because his contribution is itself significant.[17] Indeed, he was in many ways the first to

17. The introduction of the space variable into price theory was attempted much earlier than Hoover, actually more than three centuries ago by Sir William Petty, as Marx [26]

consider the possibility of price discrimination taking place under conditions of competition as well as monopoly.

Hoover assumed that marginal costs of production are constant *(c)*, that the freight rate t is an increasing function of distance, and that "all buyers have identical demand schedules," with a seller being "committed to a single grade and kind of a particular commodity."[18] He then set forth the formula recorded below as (3-1):[19]

$$p - t = \frac{ec + t}{e - 1}. \tag{3-1}$$

where p stands for delivered price, hence $p - t$ for mill price, and e for elasticity of demand. (Note that Hoover's t is equivalent to our tK of Chapter 2.)

Equation (3-1) defines the f.o.b. mill price, i.e., $p - t$, as a function of the freight rate, among other forces. It signifies that different prices are chargeable to buyers who differ in their locations, *ceteris paribus.* In other words, equation (3-1) indicates that price discrimination is possible even if every buyer has an identical demand curve, provided further that the elasticity of demand is finite and greater than unity. Of course, Hoover is talking about the third degree of discrimination, where a seller visualizes different effective demand curves for his product. Thus Hoover showed that although there may be an identical demand curve for each buyer's gross demand (i.e., the demand at the buyer's location), the net demand curves (i.e., the demand at the seller's location) are not identical, given varying spatial (distance) costs.[20] It is further the case that when buyers *are* spatially separated, resale possibilities are limited and hence help determine the direction of the discrimination. The distance or freight rate variable is thus a practically sufficient (enabling) condition per se for price discrimination.[21]

---

noted. Petty's inclusion of space was essentially confined, however, to the concept of differential rents attributable not only to differences in the fertility of lands of equal area but also to differences in location – specifically, their distance from the market.

18. Hoover [18, pp. 182–83]. Also see Smithies [34, p. 65], where similar assumptions are presented.

19. Marginal revenue, net of the freight rate, can be equated with the marginal cost of production, as Hoover [18, p. 183] suggested. Thus, $\text{MR} - t = c$, where $\text{MR} = p(1 - 1/e)$. Equation (3-1) is easily derivable from this beginning relation.

20. In fact, Hoover [18, p. 185] emphasized that the demand of a distant buyer, as viewed by the seller, is shifted downward by an amount equal to the cost of transportation from the seller's to the buyer's location.

21. Hoover [18, p. 184] notes that in actual cases the direction of the price discrimination is against nearer buyers, such as the pricing pattern of an oil company [18, pp. 187–88]. He claims theoretical support for this besides empirical support based on resale possibilities [18, p. 186]. Manifestly, discrimination against distant buyers *could* make possible resale by nearer buyers of the seller's product, and hence such discrimination will rarely be found in practice.

H. W. Singer

Shortly after Hoover's paper on spatial price discrimination, Singer [33] attempted to set forth a measure of the degree of geographical distortion caused by discriminatory pricing. His measure was designed to be analogous to measures of monopolistic distortion from the perfect market. Unfortunately, Singer's attempt is based upon complete misunderstanding of Hoover. Let us see why this is so.

Singer claimed that Hoover had shown that sellers would discriminate *against* buyers located at greater distances from them *in favor* of nearer buyers, and that this would be practiced even though buyers have "identical demand schedules of constant as well as equal elasticity" [33, p. 75]. Hoover did, in fact, *initially* assume constant and equal elasticity. But, as implied above, Hoover's theory is not at all restricted to this special case. Moreover, as has already been observed, Hoover claimed that a seller typically discriminates *against more proximate buyers* in favor of distant buyers, and he set forth the precise conditions which support his theory of discrimination. But let us probe deeper into Singer's analysis.

Singer assumed a constant (and identical) elasticity of demand for all buyers. He defined the degree of geographical distortion, call it D, as the quotient of the price charged to the "average" buyer minus the price charged the buyer at the seller's door divided by the price charged to the "average" buyer. Formula (3-2) thus follows as a matter of definition:

$$D = \frac{\dfrac{ec + T/2}{e-1} - \dfrac{ec}{e-1}}{\dfrac{ec + T/2}{e-1}}, \tag{3-2}$$

where $T/2$ is the cost of transporting a unit of the commodity to the average buyer (that is, the buyer located halfway between seller's location and his market extremity [33, p. 75]).[22] The first term in the numerator (and hence the denominator) is obtained by substituting $T/2 = t$ into Hoover's formula (3-1); hence the first term gives the mill price that is paid by the buyer located halfway distant from the seller's door to the most distant buyer. The second term in the numerator is obtained by substituting $t = 0$ into Hoover's formula (3-1), and hence stands for the mill price that is charged to a buyer located at the seller's site. The denominator completes Singer's definition of geographical distortion.

Elasticity e in the first term of the numerator of (3-2) equals e in the

22. Singer's market boundary is an exogenously given boundary or natural boundary beyond which there is no buyer. This desert, in effect, is required by the assumption of constant elasticity, for otherwise the assumption suggests that any buyer in any remote place could be reached by the seller notwithstanding high freight costs.

second term under Singer's assumption of constant elasticity. Therefore the first term in the numerator is necessarily greater than the second term, and the mill price paid by a distant buyer is necessarily higher than the mill price paid by a buyer located next to the seller. Thus, Singer is referring to price discrimination in favor of proximate buyers, precisely the case which Hoover ruled out, in part because of resale possibilities. Singer's formula thus proposes to measure the degree by which distant buyers are discriminated against in favor of nearer buyers. He opted, in other words, for a measure of spatial price discrimination which unfortunately is theoretically inconceivable and probably more often than not unobservable in practice.[23]

Leontief, Smithies, and Dewey

Spatial price discrimination may be expressed in terms of the freight absorption which characterizes the discrimination against nearer buyers [18, p. 185]. One may, in fact, describe the *rate of freight absorption as the measure of the degree* of spatial price discrimination prevailing in the system. This kind of specification was first formulated by Hoover, then followed (independently) by Leontief [20], and subsequently by Smithies [35], Dewey [5], and Greenhut [15]. In particular, Leontief proposed the existence of a *limited* discrimination compared to *unlimited* discrimination. He stated [20, p. 491] that the transportation costs which help distinguish two markets constitute "the upper limit of interregional discrimination." Significantly, this concept of *limited* discrimination tends to rule out discrimination against distant buyers in favor of proximate buyers.

Smithies sought to disprove "the widespread impression" that the spatial monopolist necessarily practices price discrimination [35, p. 73]. However, as Dewey [5, p. 50] pointed out, *what he really did was to make the monopolist's price policy depend on the shape of the demand curve.* His basic conclusions, however, had already been anticipated and given by Hoover, and further discussion of Smithies' analysis is therefore reserved for Chapter 5.

Aside from the question of who said it first, Hoover-Leontief-Smithies agreed that a spatial monopolist may or may not pursue discriminatory pricing, depending upon the shape of demand curves.[24] This conclusion

23. Singer's degree of geographical distortion is always positive because mill price for buyers located at the seller's door is lower than the mill price paid by the more distant buyers, this by his assumption of constant elasticity. The relations between elasticity and spatial prices will be discussed in connection with the rate of freight absorption in Chapter 5. We will find there a significantly different set of conclusions than that presented by Professor Singer.

24. This statement conforms to the classical, e.g. Robinsonian, proposition that discrimination depends upon the relative shapes of demand curves, i.e., upon elasticity dif-

is logically valid, but it is based implicitly on a *most unrealistic assumption,* as will also be emphasized in Chapter 5. Suffice it to say for the present that none of the aforementioned writers restricted their demand function, $p = p(q)$, such that $p(0) < \infty$ and $p^{-1}(0) < \infty$.

Dewey—in a similar key [5]—proposed that freight absorption could be related to the shapes of the average and marginal revenue curves. In a sense he argued against the Hoover-Leontief-Smithies version of the important kinds of spatial price discrimination by proposing that a negative freight absorption—phantom freight, in his words—is possible under spatial monopoly. For present purposes, let us simply suggest that phantom freight could prevail *in noncollusive systems* only if buyers do not try to profit by reselling to points where the firm is exacting a phantom freight.[25] Moreover, the demand curves which induce the firm to exact Dewey's kind of phantom freight will be shown to be highly unrealistic.

There are other types of spatial pricing techniques than the two main kinds discussed above.[26] But the emphasis later in this book on competitive f.o.b. mill pricing vis-à-vis competitive (i.e., unorganized or nonsystematic) spatial price discrimination strongly supports the importance of the simple f.o.b. and uni-directed spatial discriminatory price techniques as *the* main alternative pricing techniques available to sellers in the space economy. These pricing systems serve, moreover, as the spatial counterparts of the simple monopoly and discriminatory pricing practices which marked all of classical theory. This belief was noted years ago by Greenhut [15], who also pointed out the fundamental differences between uniform f.o.b. mill pricing (the *spatial* variant of simple monopoly pricing) and *spatial* discriminatory pricing.[27] In particular, as was noted in Chapter 2, freight absorption exists even under the f.o.b. system in the sense that the profit-maximizing f.o.b. mill price is lowered whenever the spatial monopolist expands his sales radius [15, app. A]. Apart from this characteristic of both nondiscriminatory and discriminatory pricing in economic space, the classical differences which distinguish the two kinds of pricing techniques will also be seen to hold. Discrimination by a spatial monopolist thus differs from its f.o.b. alternative in the

---

ferentials. Most particularly, observe that in economic space the relative shapes at given prices on the several *net* demand curves depend on the shape of the gross demand curves (typically assumed to be identical in spatial economics for all buyers or submarkets).

25. Collusive spatial price systems, such as the basing-point system, generate a different kind of phantom freight than that considered by Dewey.

26. Important alternative pricing techniques are the uniform delivered-price pricing system over space, and base-point pricing. The former is a special type of discriminatory pricing and the latter is a special (collusive) form of either a straight f.o.b. or a discriminatory pricing technique. Base-point pricing has drawn the attention of many economists—Fetter [13], Stocking [38], Smithies [36], Machlup [24], Kaysen [19], J. M. Clark [3], among others. And see Beckmann [1] for a brief discussion of each pricing technique.

27. See Greenhut [15, pp. 156–61], where a concise summary of some fundamental aspects of spatial price discrimination is set forth.

same general sense that price discrimination by the classical monopolist differs from simple monopoly pricing.

IV. *Preview of Chapters 4, 5, and 6*

In the next three chapters, the theories of spatial price discrimination by Hoover and others sketched above will be synthesized in such manner that a basic modification in "approach and conclusions" will be seen to be necessary. In particular, demonstration will be provided that (i) discriminatory pricing yields greater outputs than does f.o.b. pricing, *regardless of the shape of the gross demand curve,* and (ii) the spatial firm generally absorbs freight *and discriminates in price, regardless of the shape of the gross demand curve.* The conclusions to be drawn, therefore, deviate substantially from those of Hoover, Leontief, Smithies, and Dewey, and for that matter from Dupuit, Pigou, Robinson, and others.

Appendix: *The demand curve and discrimination*

This appendix relates in part to the discussions presented in notes 6 and 9 of Chapter 3. Another link is that it will help clarify the meaning of "degrees of discrimination."

I

Assume for the moment, as Hicks did implicitly, that Dupuit's demand curve is equivalent to the Marshallian *individual's* demand curve of nonzero elasticity. On this implicit assumption, Hicks [17, pp. 38–39] asserted that the slope of the indifference curve at R in Fig. 3-4 must be the same as the slope of the indifference curve at P, provided the marginal utility of money is constant. Consumers' surplus PR will then be represented by $p'Pn'$ in Fig. 3-1 above. Thus, Hicks is *in fact* asserting that

$$\frac{U_x(X_0, M_0)}{U_m(X_0, M_0)} = \frac{U_x(X_0, M_1)}{U_m(X_0, M_1)} \quad \text{if}$$

$$U_m(X_0, M_0) = U_m(X_0, M_1),$$

where U_x and U_m stand for the marginal utility of X and M (money) respectively. However, the validity of this assertion (which is basically the same as Marshall's [25, p. 106]) also requires U_x to be independent of

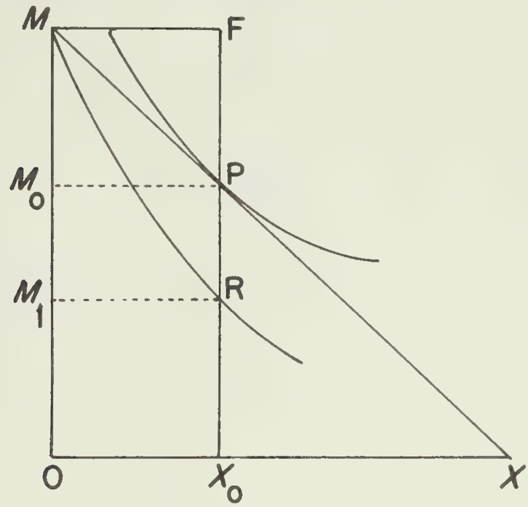


Fig. 3-4 Consumer surplus and remaining utility.

money balances. Thus, we must also have $U_x(X_0, M_0) = U_x(X_0, M_1)$; alternatively formed, $U_m(X_0, M_0) = U_m(X_1, M_0)$, and hence $U_{xm} = \partial U_x/\partial M = 0$. Only then would Marshall's "consumers' surplus" under a constant marginal utility of money be equivalent to the "utility remaining to consumers." This requirement obtains because "consumers' surplus" PR can be represented by a specific portion of the area formed by a demand curve only if the marginal rates of substitution are equal at P and R.

To understand this statement fully, let the lower and higher indifference curves in Fig. 3-4 be defined respectively as

$$M = f_1(X), \quad \text{and} \tag{3-3}$$
$$M = f_2(X). \tag{3-4}$$

Now the demand price which would provide the consumer with only the same basic "well-being" is given by $-f_{1x}(X)$, the marginal rate of substitution on the (no-better-off) indifference curve $f_1(X)$. If the equilibrium price equals $-f_{2x}(X_0)$, the difference between $-f_{1x}(X)$ and $-f_{2x}(X_0)$, when *integrated to the point where the difference vanishes*, provides the utility remaining to consumers. Taking into consideration this relationship, observe the integration:

$$\int_0^{X_0} [f_{2x}(X_0) - f_{1x}]dX = X_0 f_{2x}(X_0) - f_1(X_0) + f_1(0)$$
$$= \mathrm{M_0P(-M_0M/M_0P) - OM_1 + OM}$$
$$= \mathrm{-FP - RX_0 + FX_0 = PR.}$$

This integration, which yields PR, does not necessarily require $f_{2x}(X_0) - f_{1x}(X) = 0$ at $X = X_0$, notwithstanding the requirement that the area $Pp'n'$ in Fig. (3-1) involves integration up to the point where $f_{2x}(X_0) - f_{1x} = 0$. It follows that the area $Pp'n'$ equals PR only if $f_{2x}(X_0) = f_{1x}$ at $X = X_0$.

It must, however, be pointed out that Dupuit's demand function is not really equivalent to a Marshallian *individual* demand curve [25, p. 81], since Dupuit's demand function is actually a specific type of *market* demand curve.[28] It is special in the sense that it assumes each buyer demands only one unit of the commodity or service, an assumption not necessarily unrealistic, since, for example, no one will demand a surgical operation twice just because it is cheap.[29] Dupuit's demand curve for an individual (i.e., the demand curve comparable to the Marshallian individual demand curve) is, in other words, perfectly inelastic. Marshall's "consumers' surplus" *is represented by a portion of the area formed by a curve*, and hence is not matched by a comparable area of Dupuit's individual demand curve.[30]

The arguments above might seem to have little to do with our main subject of price discrimination. But they are not mere digressions. Rather, their incidence will be seen to lie in clarifying the meaning of Pigou's "degrees of discrimination."

II

Pigou, let us recall, referred to three degrees of discrimination. He conceived of the first degree to involve the charging of a different price for *each unit of the commodity* sold such that the price for each unit exacted equals its demand price. As a consequence, no consumers' surplus at all is left to any buyer. The second degree of discrimination is simply an incomplete form of the first degree; it involves a different price for different *quantity sets of the commodity* in such manner that some con-

28. His demand curve is not even rigorously derived from utility theory. Indeed, Walras pointed out that Dupuit's discussion of utility was vague and confused. See Walras [40, pp. 443–46]. This shortcoming does not, of course, negate the importance of Dupuit's contribution to the theory of price discrimination.

29. Dupuit's own examples are tolls on a bridge [7, pp. 107–9; 8, pp. 7–16] and railway tariffs [8, pp. 2–31]. And also see Edgeworth's discussion on railway fares [10, p. 174].

30. Note that Marshall's "consumers' surplus" per se is not based on the assumption $U_{mm} = U_{mx} = 0$ and thus is not necessarily represented by any specific portion of the demand curve area. In effect, "consumers' surplus" is defined as the excess of price which a consumer would be willing to pay over that which he actually does pay rather than do without the good (Marshall [25, p. 103]). According to this definition, "consumers' surplus" can be represented by a part of the length of the demand curve *provided it is perfectly inelastic*. Otherwise, its identification is not possible unless simplifying assumptions such as $U_{mm} = U_{mx} = 0$ are made.

sumers' surplus may be left over. The third degree would obtain if the monopolist was able to classify his *customers* by markets and could charge a different price in each [27, p. 279].

To appreciate the meaning of the first degree of discrimination, we paraphrase Due [6, p. 250]. He stated that if a consumer would purchase 1 gallon of gasoline each week at $1.00 a gallon, 5 gallons at 50 cents a gallon, and 10 gallons if the price was 25 cents a gallon, then perfect discrimination would involve charging him $1.00 for one gallon, 50 cents for the next 4 gallons, and then 25 cents for the last 5. But this statement violates a fundamental property of the theory of demand, because the demand schedule of the buyer cannot in general remain the same if he has already committed himself to buy 1 gallon of gasoline; i.e., his money balance is now lower than before. Demand for gasoline is fundamentally a function of real balances, *the* amount of gasoline the buyer has obtained, the prevailing price, and other forces [16, chap. 2].

The point at issue may be illustrated diagrammatically in Fig. 3-5. The horizontal axis is designed to measure the amount of gasoline, while the vertical axis measures the monetary balance. Assume the initial equilibrium point is at P, given the budget constraint shown in panel (a) of Fig. 3-5. According to Due, a price reduction would yield the new equilibrium value at Q. Given this change in price, the buyer would purchase 4 extra gallons of gasoline. But the true solution point is at R, where the budget line PR (parallel to MQ) is tangent to the indifference

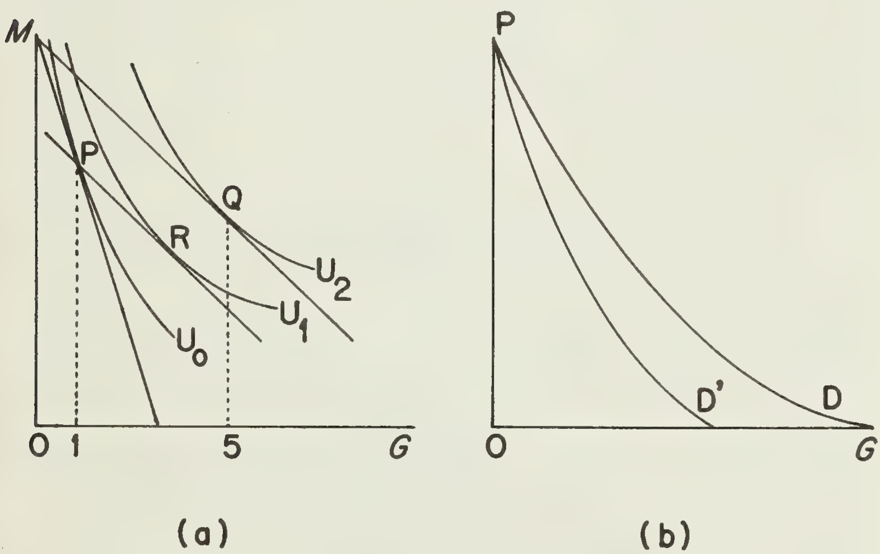


Fig. 3-5 The meaning of first-degree discrimination.

curve U_1. This solution assumes that both gasoline and money are normal goods, and accordingly it yields less extra demand for gasoline than that shown by Q when price is reduced. The price-quantity relationship under Pigou's first degree of discrimination is illustrated by the PD′ curve, given the Marshallian demand curve PD.[31] The two curves merge into PD if the marginal utility of money is both constant and independent of the amount of gasoline the buyer has obtained.[32] If the marginal utility of money is not independent of gasoline and increases as more gasoline is purchased, the real demand curve would be PD′.[33]

Mrs. Robinson [28, p. 187] argued that Pigou's first degree of discrimination would obtain only if each consumer bought only one unit of the product. But this is not a necessary condition, as Due's example demonstrates. In particular, a buyer may purchase more than one unit of product under Pigou's first degree of discrimination. However, the first degree of discrimination in Due's example does presuppose (i.e., require) that the related third degree of discrimination would involve complete separation of *all* buyers and complete appropriation by the seller of each buyer's "consumers' surplus."

In passing, observe that Dupuit's discrimination is a special case in which Pigou's first and third degree of discrimination occur simultaneously as above. This is so because a complete separation of buyers implies and is implied by the corresponding complete separation of commodity (service) units.[34] It follows that if a seller can classify each and every consumer according to his effective demand for the good and charge each a different price, the limiting case obtains where Pigou's third degree of discrimination converges to the first.[35] And we see that if each buyer is charged the maximum price he would be prepared to pay for a good rather than go without it, the situation involves nothing but a complete separation of commodity units sold with no consumers' surplus left, i.e., discrimination of the first degree. Moreover, as with Dupuit,

31. PD is the Marshallian demand curve while the curve PD′ is equivalent to Slutsky's constant-apparent-real-income demand curve (see Slutsky [34, pp. 466–68] or Friedman [14, p. 51]). Hicks's interpretation of the Marshallian demand curve would yield a PD curve. Due, in turn, points to the Marshallian demand curve PD. (Note that PD implicitly yields increasing income, i.e., consumer satisfactions, at lower prices and therefore violates the *ceteris paribus* pound so fundamental to demand curve theory; see [14].)

32. Let it be recalled that these two conditions are implicitly required by the concept that "consumers' surplus" is representable by a portion of the area formed by a demand curve. Needless to say, neither the Marshallian demand curve per se nor Marshall's analysis of "consumers' surplus" requires these conditions.

33. Significantly, each additional unit sold under the first degree of discrimination adds to revenue an amount equal to the price for which it is sold. Thus, the real demand or demand locus PD′ (not PD) is also the marginal revenue curve of the monopolist.

34. See note 9 in this chapter.

35. This is the limiting case because each group of customers or each submarket contains only one buyer by assumption.

if the individual demand curve is so inelastic that each person purchases either one unit of the commodity (or service) or none, depending upon the price level, then the discrimination would be of the order of both the first and third degrees, since each and every buyer *and* each and every unit of the good are separable. Our conclusion is that the first (or second) and third degrees of discrimination are distinguishable only under a model where individual demand curves have finite nonzero elasticity; they are not distinctive under Dupuit's model where a perfectly inelastic demand curve is assumed.

References

[1] M. Beckmann. *Location Theory.* New York, 1968.
[2] E. H. Chamberlin. *The Theory of Monopolistic Competition.* Cambridge, Mass., 1946.
[3] J. M. Clark. "Basing Point Methods of Price Quoting." *Canadian Journal of Economics and Political Science,* 1938, 4:477–89.
[4] A. A. Cournot. *Researches Into the Mathematical Principles of the Theory of Wealth.* Translated by N. T. Bacon. 1897.
[5] D. Dewey. "A Reappraisal of F.O.B. Pricing and Freight Absorption." *Southern Economic Journal,* July 1955, 22:48–54.
[6] J. F. Due. *Intermediate Economic Analysis.* Homewood, Ill., 1956.
[7] J. Dupuit. "On the Measurement of the Utility of Public Works." *International Economic Papers* (London, 1952), pp. 83–110.
[8] J. Dupuit. "On Tolls and Transport Charges," *International Economic Papers* (London, 1962), pp. 7–31.
[9] F. Y. Edgeworth. "Applications of Probabilities to Economics." *Economic Journal,* 1910, 20:284–314 and 441–65.
[10] F. Y. Edgeworth. *Papers Relating to Political Economy.* London, 1925.
[11] E. O. Edwards. "The Analysis of Output Under Discrimination." *Econometrica,* 1950, 18:168–72.
[12] R. B. Ekelund Jr. "Price Discrimination and Product Differentiation in *Economic Theory:* An Early Analysis." *Quarterly Journal of Economics,* 1970, 84:268–78.
[13] F. A. Fetter. "The New Plea for Basing Point Monopoly." *Journal of Political Economy,* 1937, 45:577–605.
[14] M. Friedman. "The Marshallian Demand Curve." *Journal of Political Economy,* 1949, 57:463–95.
[15] M. L. Greenhut. *Plant Location in Theory and in Practice.* Chapel Hill, N.C., 1956.
[16] M. L. Greenhut. *Microeconomics and the Space Economy.* Chicago, 1963.
[17] J. R. Hicks. *Value and Capital.* 2d edition. London, 1946.
[18] E. M. Hoover. "Spatial Price Discrimination." *Review of Economic Studies,* 1937, 4:182–91.
[19] C. Kaysen. "Basing Point Pricing and Public Policy." *Quarterly Journal of Economics,* 1949, 63:289–314.
[20] W. Leontief. "The Theory of Limited and Unlimited Discrimination." *Quarterly Journal of Economics,* 1939, 54:490–501.
[21] A. P. Lerner. "The Concept of Monopoly and the Measurement of Monopoly Power." *Review of Economic Studies,* 1933, 1:157–75.
[22] A. P. Lerner and H. W. Singer. "Some Notes on Duopoly and Spatial Competition." *Journal of Political Economy,* 1937, 45:145–86.
[23] A. Lösch. *The Economics of Location.* New Haven, 1954.
[24] F. Machlup. *The Basing Point System.* Philadelphia, 1949.
[25] A. Marshall. *Principles of Economics.* 8th edition. London, 1961.

[26] K. Marx. *A History of Economic Theories.* Translated by T. McCarthy. New York, 1952.
[27] Y. R. Moroni. "Discrimination Under Market Interdependence." *Quarterly Journal of Economics*, 1947, 61:95–117.
[28] A. C. Pigou. *The Economics of Welfare.* 3d edition. London, 1929.
[29] J. Robinson. *The Economics of Imperfect Competition.* London, 1933.
[30] P. A. Samuelson. *Foundations of Economic Analysis.* Cambridge, Mass., 1947.
[31] J. A. Schumpeter. *History of Economic Analysis.* New York, 1954.
[32] C. G. F. Simkin. "Some Aspects and Generalization of the Theory of Discrimination." *Review of Economic Studies*, 1948/49, 15:1–13.
[33] H. W. Singer. "A Note on Spatial Price Discrimination," *Review of Economic Studies*, Oct. 1937, 5:75–77.
[34] Eugenio Slutsky. "Sulla teoria del bilancio del consumatore." *Giornale degli Economisti*, 1915, 51:1–26, esp. sec. 8.
[35] A. Smithies. "Monopolistic Price Policy in a Spatial Market." *Econometrica*, 1941, 9:63–73.
[36] A. Smithies. "Aspects of the Basing Point System." *American Economic Review*, 1942, 32:705–26.
[37] G. Stigler. *The Theory of Price.* New York, 1966.
[38] W. Stocking. *Basing Point Pricing and Regional Development.* Chapel Hill, N.C., 1954.
[39] J. Viner. *Dumping, a Problem in International Trade.* New York, 1966.
[40] L. Walras. *Elements of Pure Economics.* Translated by W. Jaffé. London, 1954.
[41] T. D. Yntema. "The Influence of Dumping on Monopoly Price." *Journal of Political Economy*, 1928, 36:686–98.

4. A theory of nondiscriminatory and discriminatory spatial monopoly pricing

I. *Introduction*

This chapter is concerned with certain special characteristics of spatial price discrimination. Following initially along the lines of [2], it establishes the proposition that spatial price discrimination always yields greater outputs than does f.o.b. mill pricing. This result is shown to apply *regardless of the shape of the gross demand curve,* where by gross demand curve we mean the buyer's demand for a product if the cost of distance was zero. As a result of the present inquiry, Chapter 5 will then be in position to synthesize the Hoover, Leontief, Smithies, Dewey approaches to spatial price discrimination and to establish directly the related propositions that not only is spatial price discrimination more profitable than f.o.b. mill pricing, but it is always directed against the buyers located nearest to the seller, *again regardless of the shape of the gross demand curve.*

II. *A basic difference in nonspatial and spatial price discrimination: total output*

The discriminating monopolist is described in classical (nonspatial) economics as the supplier of greater *or* smaller outputs than the nondiscriminating monopolist, depending upon the relative shapes of the demand curves. However, as was shown in Chapter 3, whether or not more outputs are produced under discriminatory monopoly also depends in part on the level and/or shape of the cost function. It is this extra property which is the relevant force, in fact the basic force, behind the alternative views that apply to nonspatial price discrimination. The purpose of the present chapter is not to take issue with classical theory, but to reformulate it so that the pure implication of economic space can be evaluated. Apart from discrimination by licensed practitioners of medi-

cine, law, and dentistry, much of the price discrimination by firms encountered today is structured along geographical lines.

The spaceless framework

As the departure point for our basic model, assume as did Pigou [3] and Mrs. Robinson [4] the existence of two markets, each of which possesses a linear demand curve. Assume these curves are parallel and differ only in price and quantity intercepts. Let market I contain the stronger demand while market II is characterized by the weaker demand. Assume that there is no cost of distance, or alternatively that the economic spaces separating the markets from the seller are the same.

Fig. 4-1(a) and (b) respectively measure the individual demand q_i and the aggregate demand Q of the two markets on the horizontal axes. The vertical axes measure average revenue as well as marginal revenue for the simple and discriminatory monopolies. Lines p_1q_1 and p_2q_2 in Fig. 4-1(a) stand for the demand curve respectively in market I and market II. The line pRQ in Fig. 4-1(b) is the horizontal sum of the two demand curves. This aggregate demand curve (AD) applies to a simple monopoly. Its marginal revenue curve SMR is the broken curve *pbcd*.[1] Correspondingly, the aggregate *demand* curve under discriminatory monopoly is also AD; but, an average *revenue* curve is conceptually imaginable which – in view of the different prices charged in the two markets – would yield for the discriminating monopolist a higher net price than that of the simple monopolist when sales are made to both markets and total outputs are identical. (We do not show this average revenue curve in Fig. 4-1). The *continuous curve pacd* is the discriminatory marginal revenue curve DMR. It is the horizontal sum of the two marginal revenue curves and must be distinguished from the discontinuous SMR curve *pbcd*. The intersection of DMR and SMR with the marginal cost curve MC determines whether or not total output Q is greater under discriminatory monopoly than simple monopoly.

Case I. If MC cuts the *pa* portion of DMR, there will be no difference between the outputs produced by a discriminatory monopoly and a simple monopoly. In effect, only one market – market I – is served *even if the monopolist otherwise would practice discrimination* (see J. Robinson [4, p. 196]).

Case II. If MC cuts the *ab* portion of SMR, and also the *ac* portion of DMR, the intersection of MC with DMR lies to the right of the inter-

1. This diagram, in particular Fig. 4-1(b), is based on J. Robinson's Fig. 65A [4, p. 201]. A diagram by Pigou [3, p. 809], designed for similar analytical purposes, is inferior to Robinson's in the sense that it requires a constant marginal cost (Pigou's supply price).

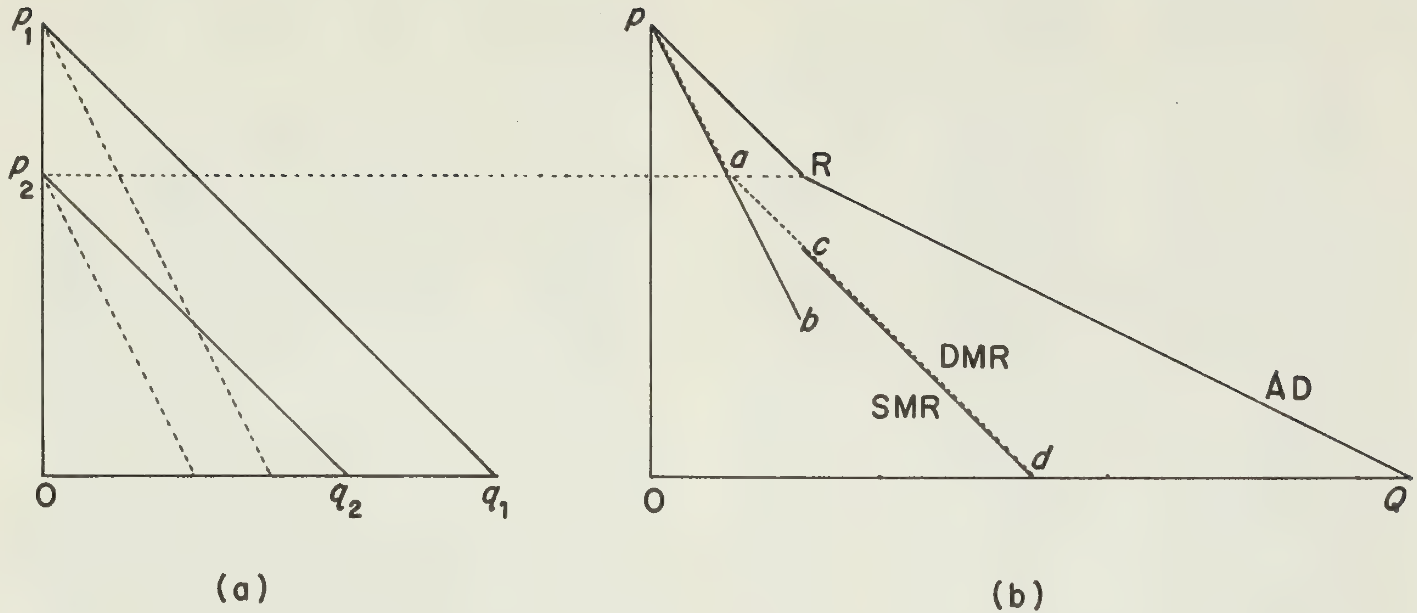


Fig. 4-1 Aggregate demand and alternative marginal revenue curves: the nonspatial case.

section of MC with SMR. Total output is, therefore, greater under discriminatory pricing than under simple monopoly pricing (and see A. C. Pigou [3, pp. 286, 809]). Market I alone is served in this case by the simple monopolist whereas both markets are served by the discriminating seller. Correspondingly, the equilibrium price is necessarily higher in the case of simple monopoly pricing than p_2 while, in contrast, the price in market II is lower than p_2 in the price discrimination case (Fig. 4-1).

Case III. If MC cuts the *cd* portion of DMR (which is overlapped by SMR), total output is once again the same under the two pricing systems. Unlike Case I or II, however, both markets are now served. Although total output remains unchanged, the output shipped to market II increases by the precise amount by which demand is cut in market I as the pricing technique is changed from simple to discriminatory pricing.[2]

Of the three cases, Mrs. Robinson apparently emphasized Case III. She identified it with the rather elementary situation of linear demand [4, p. 192]. Although Case I may be trivial, Case II will be shown to be highly relevant to a spatial analysis of discriminatory pricing.

The spatial framework

To appreciate the relationships relevant to the space economy, assume for simplicity that buyers are *evenly distributed* along a line or over a plain and that every buyer has the same identical (gross) demand curve.[3] Then the net demand curve of a buyer (or group of buyers) one mile away from the seller must be different (to the seller) than the demand curve of a buyer located at the seller's door. The net demand curves of a buyer two miles away, a buyer three miles away, etc. are in turn all different in the eyes of the seller. A number of parallel linear functions can thus be used to portray the net demands confronting the seller. While a single seller cannot serve spatial markets in which the freight cost burden is prohibitive, it should be clear that more distant buyers (or markets) may be served by a discriminating monopolist compared to a simple f.o.b. pricing monopolist. Cases I and III above are, therefore, inapplicable, and Case II alone is relevant under spatial monopoly. But let us probe into this matter more deeply as follows.

Consider three different net demand curves belonging, respectively, to markets located adjacent to, near to, and rather distant from the seller.

2. Pigou [3, p. 809]; Robinson [4, p. 192]; and see Battalio and Ekelund [1] for a lucid discussion on the importance of cost in price discrimination.

3. This spatial system conforms to the Hoover-Smithies-Dewey framework of thought (see Chapter 5 for details).

Assume further that: (1) Every buyer's gross demand curve is linear and identical of the form $p = b - 3aq$ so that the aggregate gross demand curve is specifiable as $p = b - aQ$, where b and a are positive constants. (2) The freight rate (or cost of distance) is zero between the seller and the adjacent market, $(\frac{1}{3})b$ over the distance from the seller to the near market, and $(\frac{2}{3})b$ for the distance between the seller and the distant market. (3) Each market contains $\frac{1}{3}$ of the total number of buyers. Fig. 4-2 may, then, be constructed as a simple extension of Fig. 4-1. However, since we are here conceiving of three *net* demand curves, i.e., in terms of net *mill* prices, the vertical axis may now be represented by m instead of p. Thus, for example, in Fig. 4-2(a) $m_i q_i$ are depicted to represent net individual demand curves while mRSQ in panel (b) stands for aggregate net demand. This aggregate net demand is specifiable algebraically by

$$\begin{aligned} Q &= \frac{b-m}{3a}, \quad \forall m \ni m_1 > m \geqslant m_2 \\ &= \frac{b-m}{3a} + \frac{b-(b/3)-m}{3a}, \quad \forall m \ni m_2 > m \geqslant m_3 \\ &= \frac{b-m}{3a} + \frac{b-(b/3)-m}{3a} + \frac{b-(2b/3)-m}{3a}, \quad \forall m \ni m_3 > m \geqslant 0. \end{aligned}$$

The aggregation of individual demands illustrated above is counterpart to the kinked demand curve mRSQ. Correspondingly, the DMR and SMR curves are rigorously specifiable for the subject case; however, this specification will be deferred to the generalized formulae derived later as (4-2) and (4-3).

Observe that Case III, i.e., where DMR and SMR merge into the sequent *cd* in Fig. 4-1, is now relatively unimportant since the line *cd* has become relatively shorter in a threefold market division vis-à-vis the old twofold division of buyers. A new *df* portion of DMR has appeared which is *not overlapped* by SMR. One might well expect, accordingly, that outputs will often be different.

The following conclusions apply: If marginal cost is so high that the MC curve cuts the *ac* portion of DMR, total output will clearly be greater under discrimination than it would be under simple monopoly. If the MC curve cuts the *cd* portion of DMR = SMR, the total outputs will be the same. If MC is still lower so that it cuts the *df* portion of DMR, total output once again will be greater under discriminatory pricing than under simple monopoly pricing. Diverging total outputs become more and more likely under linear demand the "finer" the division of markets in economic space.

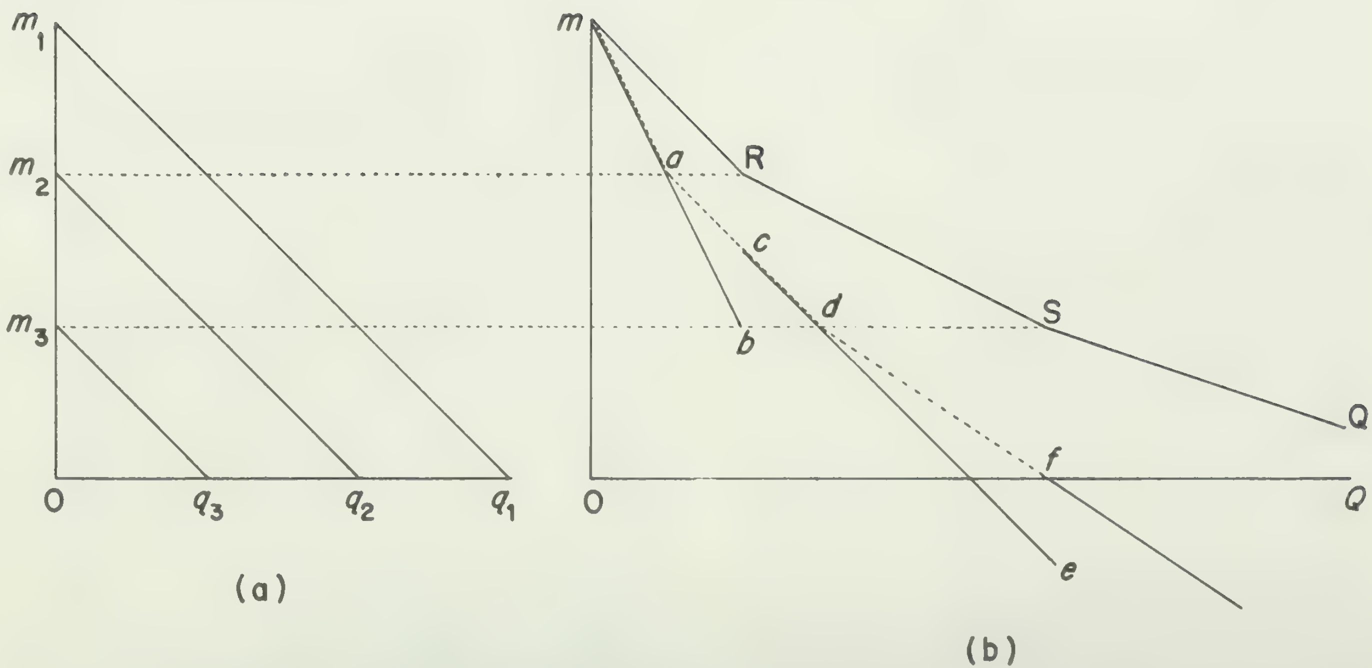


Fig. 4·2 Aggregate demand and alternative marginal revenue curves: the spatial case.

III. *The spatial model expanded and generalized*

Consider the "line" situation where n spatially separated markets of the same size are identifiable, and there is an identical gross demand curve for each submarket. The freight rate per unit of distance may now be defined by the constant value b/n, where b stands for the price intercept of the gross demand curve of a buyer at a given site on the line market. Thus, for example, the buying market one distance unit from the seller involves a freight rate equal to b/n, the buying market two distance units from the seller involves a freight rate equal to $2b/n$, etc. The price intercepts $(m_1, m_2, \ldots, m_n)$ of the net linear demand curves are, therefore, definable as $m_i = b - (i-1)b/n$, where $i = 1, 2, \ldots, n$. (See Fig. 4-3.) The spatial aggregate demand function may then be derived algebraically by summing all of the net demand curves. Thus:

$$Q = \frac{b-m}{na}, \quad \forall m \ni m_1 > m \geqslant m_2$$

$$= \frac{b-m}{na} + \frac{b-(b/n)-m}{na} = \frac{2b-(b/n)-2m}{na}, \quad \forall m \ni m_2 > m \geqslant m_3$$

$$= \frac{2b-(b/n)-2m}{na} + \frac{b-(2b/n)-m}{na} = \frac{3b-(3b/n)-3m}{na}, \quad \forall m \ni m_3 > m \geqslant m_4$$

$$= \frac{3b-(3b/n)-3m}{na} + \frac{b-(3b/n)-m}{na} = \frac{4b-(6b/n)-4m}{na}, \quad \forall m \ni m_4 > m \geqslant m_5$$

$$= \frac{i(b-m)-i(i-1)b/2n}{na}, \quad \forall m \ni m_i > m \geqslant m_{i+1}$$

where, as before, $m_i = b-(i-1)b/n$, $i = 1, 2, \ldots, n$, and n stands for the greatest number of submarkets visualized, but not necessarily served, by the firm. This background equation establishes the general formula for the spatial demand function:

$$Q = \frac{i[(b-m)-(i-1)b/2n]}{na}, \quad \forall m \ni m_i > m \geqslant m_{i+1},\ i = 1, 2, \ldots, n. \tag{4-1}$$

It is significant that as n approaches a large number, any m in the domain $m_i > m > m_{i+1}$ can be approximated by $m = (m_i + m_{i+1})/2 = [2-(2i-1)/n](b/2)$ by substitution of the price intercepts m_i. It there-

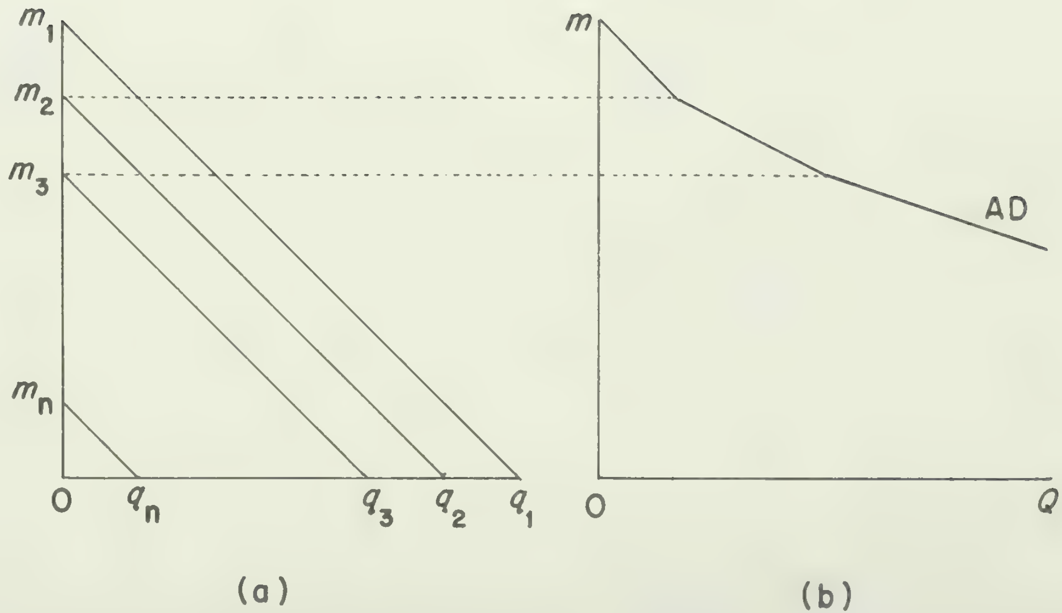


Fig. 4-3 Spatial aggregate demand in the limit.

fore follows that $i = (1/2) + (b - m)n/b$. Q may then be defined in terms of m alone by substituting this value of i into the general formula for Q; we obtain

$$Q = \frac{b-m}{2na} - \frac{b[(b-m)(n/b)-(1/2)]}{4n^2a} + \frac{(b-m)^2}{ab} - \frac{(b-m)^2(n/b)-(b-m)/2}{2na},$$

$$\lim_{n\to\infty} Q = \frac{(b-m)^2}{2ab}, \quad \forall m \ni b \geqslant m \geqslant 0, \tag{4-1$'$}$$

the same as $(f(0) - m)^2/2af(0)$ in Chapter 2, where $b = f(0)$.

Since derivation of the simple marginal revenue SMR from the spatial aggregate demand curve is somewhat more involved than that of the discriminatory marginal revenue DMR, let us begin with the latter. The DMR is easily obtained by summing horizontally all of the net individual marginal revenue curves, where each horizontal intercept value is one-half that of the respective net demand curve. Dividing the right-hand side of (4-1) by 2, therefore, yields DMR or, more rigorously, it yields the inverse function of aggregate marginal revenue under price discrimination. Thus:

$$Q = \frac{i[(b-m)-(i-1)b/2n]}{2na}, \tag{4-2}$$

$$\forall m \ni m_i > m \geqslant m_{i+1},\ i = 1, 2, \ldots, n,$$

$$\lim_{n\to\infty} Q = \frac{(b-m)^2}{4ab}, \quad \forall m \ni b \geqslant m \geqslant 0, \tag{4-2$'$}$$

where m now stands for values *to be read along the marginal revenue curve.* This formula (4-2) or (4-2)′ provides the total output produced by the seller under the specification that marginal revenue equals marginal cost.

Consider now the significance of SMR for later comparison with DMR. Observe first that it is *not* a continuous function *nor* a one-to-one function (see Fig. 4-2). In other words, even though m may be taken as a function of Q, the Q relating to SMR is not a function of m, since more than one Q value relates to some of these m's. Since SMR is not a one-to-one function, no inverse function may be derived for it that would compare with (4-2). We may, nevertheless, rewrite (4-2) as (4-3) and subsequently specify a new domain for it. In particular, we obtain

$$m = b - \frac{(i-1)}{2}\frac{b}{n} - \frac{2na}{i}Q, \quad i = 1, 2, \ldots, n, \tag{4-3}$$

where, to repeat, the domain must be redefined to establish a different function. We define this domain D by (4-3D), as explained below.[4] Thus:

$$\frac{(i+1)ib}{2n^2a} > Q \geq \frac{(i-1)ib}{2n^2a}, \quad i = 1, 2, \ldots, n. \tag{4-3D}$$

Formulas (4-3D) and (4-3) therefore respectively establish the domain and the corresponding range of the marginal revenue of the simple (nondiscriminatory) monopolist.

Note that the spatial aggregate demand AD, the discriminatory marginal revenue DMR, and the simple marginal revenue SMR are now completely specifiable via (4-1) through (4-3) and (4-3D), once the number of visualized submarkets n is known. Fig. 4-4 illustrates these curves for the case when $n = 6$. It should be emphasized that these formulae and a figure supported by them are generally applicable for the case where buyers are discretely distributed along a line market, i.e., when n is finite. This is a remarkable result! Our analysis is not confined to the conventional limiting case of continuous buyer distribution. For completeness, however, let us consider the limiting case too, the formulas for which were already given by (4-1)′ and (4-2)′; thus, there remains simply the need to derive the limiting formula for (4-3).

Profit maximization under simple monopoly requires the level of m in (4-3), i.e., SMR, to be equated with the level of MC. Hence the total output produced, given the MC, can be obtained by (4-3) if i is also given. And i (the number of markets served) turns out to be a function of m as n approaches a large number. By elementary substitutions, the equation $i/n = 2(b - m)/3b$ is obtained.[(1)] Substituting this result into (4-3) and taking the limit yields

$$\lim_{n \to \infty} Q = \frac{2(b-m)^2}{9ab}, \quad \forall m \ni b \geq m \geq 0, \tag{4-3)′}$$

the same result as that given by (2-18). Thus (4-3)′ provides the total output produced by a nondiscriminating spatial monopolist as n approaches infinity.[5] In fact, it does more than this when the entire formulation leading to (4-3)′ is summarily brought into focus. In particular, it

4. The domain Q for SMR may be partitioned for each point of Q at which a kink in the aggregate demand AD occurs. Such kinks in AD occur at the values $m = b - (i-1)b/n$, $i = 1, 2, \ldots, n$. Substituting this specific value of m into (4-1) then yields the critical values of Q at which kinks of AD occur; namely, $Q_i = i(i-1)b/2n^2a$, and hence $Q_{i+1} = (i+1)ib/2n^2a$. These values respectively provide the lower and upper limits for the relevant partitioned domain given by (4-3D).

5. It might also be noted that (4-3)′ conforms to the limiting aggregate demand function (4-1)′, actually being its marginal curve.

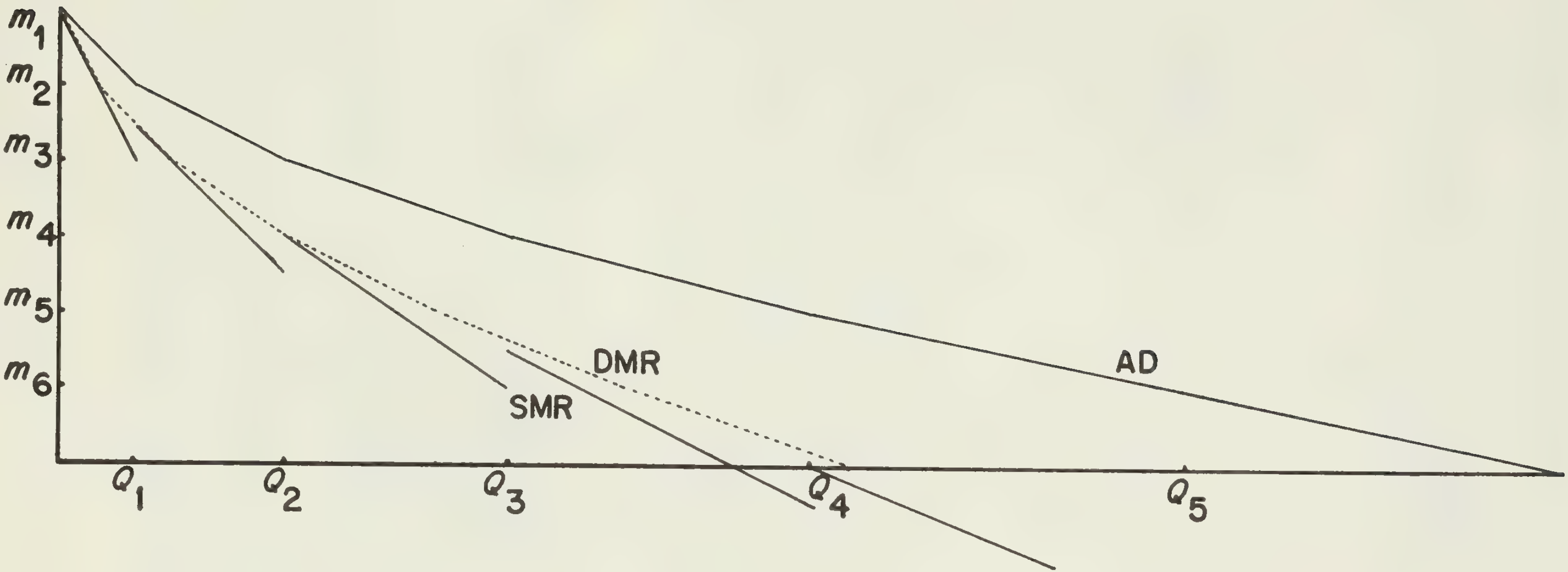


Fig. 4-4 AD, DMR, and SMR when $n = 6$.

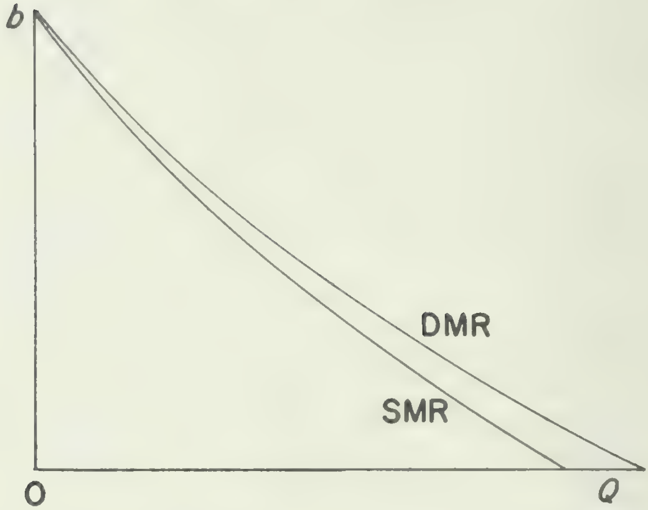


Fig. 4-5 Output is always greater under spatial price discrimination.

should be manifest that when the spatial market is "finely" divided, Q becomes a one-to-one function of m over the domain $b \geqslant m \geqslant 0$. This "fine" division also implies that the marginal revenue curve under simple monopoly, i.e., SMR, approaches a continuous curve, as shown in Fig. 4-5. But then the limiting SMR in (4-3)′ can be compared with the limiting DMR of (4-2)′. Manifestly, the value Q in (4-2)′ is strictly greater than Q in (4-3)′ *for all m such that* $b \geqslant m \geqslant 0$, and accordingly we have

$$\frac{(b-m)^2}{4ab} > \frac{2(b-m)^2}{9ab}, \quad \forall m \ni b > m \geqslant 0. \tag{4-4}$$

This relation is sufficient to support Fig. 4-5. Besides continuity, Fig. 4-5 reveals that the DMR curve always lies above the SMR curve, except at $m = b$. The quantity produced under spatial discrimination must therefore always be greater in the case of linear gross demand than that produced under simple spatial monopoly.

IV. *The spatial model generalized for nonlinear demand*

Consider once more the aggregate as well as individual gross demand function:

$$\begin{aligned} p &= f(Q) \\ &= f(nq), \end{aligned} \tag{4-5}$$

where p stands for price, Q for aggregate demand, q for individual demand, and n for the number of homogeneous individuals or submarkets. In view of some economically meaningful constraints, such as $df/dq < 0$, $0 < f(0) < \infty$, and $0 < f^{-1}(0) < \infty$, (4-5) may, in turn, be specified as

$$\begin{aligned} p &= b - aQ^{\alpha} \\ &= b - a(nq)^{\alpha}, \end{aligned} \tag{4-5$'$}$$

where b, a, and α are positive (and finite) constants. Note that the special case where $\alpha = 1$ is the linear demand case.[6]

Given the distribution of submarkets outlined in Section III above, a spatial monopolist would visualize the following n different *net* demand functions:

$$\begin{aligned} m &= f(nq) - \frac{(i-1)f(0)}{n} \\ &= b - a(nq)^{\alpha} - \frac{(i-1)b}{n}, \end{aligned} \tag{4-6}$$

or, equivalently, by rearranging terms:

$$q = \frac{1}{n}\left[\frac{b-m}{a} - \frac{(i-1)b}{an}\right]^{1/\alpha}, \quad \forall i \ni i = 1, 2, \ldots, n. \tag{4-6$'$}$$

Therefore, aggregate demand Q is obtainable via simple summation of (4-6)′. Thus:

$$Q = \frac{1}{n}\sum_{i=1}^{k}\left[\frac{b-m}{a} - \frac{(i-1)b}{an}\right]^{1/\alpha}, \qquad \forall m \ni b - \frac{(k-1)b}{n} > m \geq b - \frac{kb}{n},\ k = 1, 2, \ldots, n, \tag{4-7}$$

where k represents the most distant buying point possible given the level of mill price contemplated by the seller. Equation (4-7) can be converted to integral form as n approaches ∞. Letting $Q^* = \lim_{n\to\infty} Q$, and $x = (i-1)b/n$, so that $dx = b/n$, we obtain[II]

$$Q^* = \frac{1}{b}\int_0^{x_0}\left(\frac{b-m-x}{a}\right)^{1/\alpha} dx, \tag{4-7$'$}$$

6. The form or concavity of the demand function (4-5)′ depends upon the value of α. If $\alpha > 1$, a convex curve exists while if $\alpha < 1$, the curve is concave.

where $x_0 = (k-1)b/n$ and $k = 1, 2, \ldots, n$. Equation (4-7)′ defines the aggregate demand in the limit as a function of k, not directly as a function of m. However, since the firm's market boundary extends to the point where demand vanishes for any given m, i.e., $x_0 = b - m$, (4-7)′ reappears as

$$Q^* = \frac{1}{b}\int_0^{b-m}\left(\frac{b-m-x}{a}\right)^{1/\alpha} dx = \frac{a\alpha}{b(1+\alpha)}\left(\frac{b-m}{a}\right)^{1+1/\alpha}, \quad \forall m \ni b \geq m \geq 0. \tag{4-7}''$$

Equation (4-7)″ is a general limiting formula for aggregate demand, i.e., as $n \to \infty$.

The inverse function of simple marginal revenue (SMR) and the output of the simple monopolist may be obtained directly from (4-7)″.(III) This provides

$$Q_s^* = \frac{a\alpha}{b(1+\alpha)}\left(\frac{1+\alpha}{1+2\alpha}\right)^{1+1/\alpha}\left(\frac{b-R}{a}\right)^{1+1/\alpha}, \quad \forall R \ni b \geq R \geq 0, \tag{4-8}$$

where Q_s^* refers to output supplied under simple monopoly and where for present notational simplicity R is used now in place of m to represent the value of marginal revenue equal to marginal cost. (Note that R in (4-8) stands for SMR.)

Consider next the individual marginal revenue functions.[7] Their sum provides the discriminatory marginal revenue DMR, or its inverse function. Significantly, the individual marginal revenue functions can be derived from (4-6)′,(IV) which provides

$$R = b - a(1+\alpha)(nq)^\alpha - \frac{(i-1)b}{n}, \quad \forall i \ni i = 1, 2, \ldots, n, \tag{4-9}$$

or, equivalently, by rearranging terms,

$$q = \frac{1}{n}\left[\frac{b - R - (i-1)b/n}{a(1+\alpha)}\right]^{1/\alpha}, \quad \forall i \ni i = 1, 2, \ldots, n. \tag{4-9}'$$

The DMR, or its inverse function, involves summation of (4-9)′. Thus:

7. The general formula for the individual marginal revenue function is $R = f(nq)(1 + 1/e(q)) - (i-1)f(0)/n$, where $e(q) = (dq/dm)(m/q)$. Unfortunately, the inverse of this function is not specifiable unless f and correspondingly e are specified. Therefore, aggregation of q to obtain DMR requires specification of (4-5), e.g., as (4-5)′.

$$Q = \frac{1}{n}\sum_{i=1}^{k}\left[\frac{b - R - (i-1)b/n}{a(1+\alpha)}\right]^{1/\alpha}, \qquad (4\text{-}10)$$

$$\forall R \ni b - \frac{(k-1)b}{n} > R \geq b - \frac{kb}{n},\ k = 1, 2, \ldots, n.$$

This equation can be transformed to an integral equation as $n \to \infty$.[V] Thus:

$$Q^* = \frac{1}{b}\int_0^{b-R}\left[\frac{b-R-x}{a(1+\alpha)}\right]^{1/\alpha} dx. \qquad (4\text{-}10)'$$

Evaluating this integral[VI] gives

$$Q_d{}^* = \frac{a\alpha}{b}\left(\frac{1}{1+\alpha}\right)^{1+1/\alpha}\left(\frac{b-R}{a}\right)^{1+1/\alpha}, \quad \forall R \ni b \geq R \geq 0, \qquad (4\text{-}10)''$$

where $Q_d{}^*$ refers to output produced under discrimination.

The final steps of comparing outputs produced under alternative pricing techniques, which is equivalent to comparing $Q_d{}^*$ with $Q_s{}^*$ are at hand. For this purpose, divide equation (4-10)″ by (4-8), and record the result as

$$\frac{Q_d{}^*}{Q_s{}^*} = y^{1/\alpha}, \quad \text{where } y = \frac{1}{1+\alpha}\left(1 + \frac{\alpha}{1+\alpha}\right)^{1+\alpha}. \qquad (4\text{-}11)$$

To prove that y in (4-11) is strictly greater than unity if $\alpha > 0$, consider the following real function:

$$g(x) = x^{1+\alpha}, \quad \forall x \ni x > 0. \qquad (4\text{-}12)$$

This function implies the inequality of (4-13), as depicted in Fig. 4-6:

$$g(1+h) > g(1) + g'(1)h, \quad g' = \frac{dg}{dx}. \qquad (4\text{-}13)$$

The relation (4-13) is equivalently specifiable via (4-12) as

$$(1+h)^{1+\alpha} > 1 + (1+\alpha)h, \quad \forall\, \alpha \,\&\, h \ni h > 0 \,\&\, \alpha > 0. \qquad (4\text{-}13)'$$

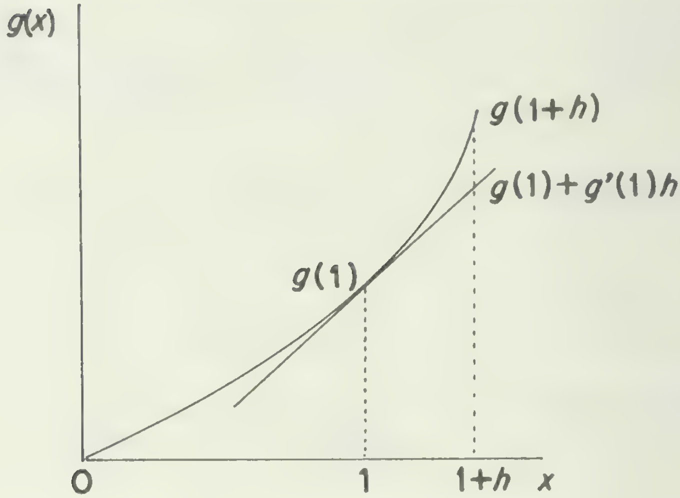


Fig. 4-6 A curve convex downward and its straight-line tangent.

Substituting $h = \alpha/(1 + \alpha)$ into (4-13)′ yields

$$\left(1 + \frac{\alpha}{1+\alpha}\right)^{1+\alpha} > 1 + \alpha, \quad \text{i.e.,} \tag{4-13)''$$

$$\frac{1}{1+\alpha}\left(1 + \frac{\alpha}{1+\alpha}\right)^{1+\alpha} > 1, \quad \forall\, \alpha > 0.$$

Thus, (4-13)″ proves that y in (4-11) is strictly greater than unity. Furthermore, since $\alpha > 0$ and $y > 1$, $y^{1/\alpha}$ must also be greater than unity.

Proof exists that Q_d^*/Q_s^* is greater than unity, i.e., Q_d^* is greater than Q_s^*, irrespective of the value of α, provided $\alpha > 0$. In other words, output produced under spatial price discrimination is necessarily greater in the limit than that produced under simple spatial monopoly regardless of the shape or concavity (convexity) of the demand function. Fig. 4-5 always applies in the space economy.

The several conclusions drawn above stem from the assumptions that buyers are evenly distributed along a line or over a plain, with each having identical tastes and hence identical gross demand curves. These assumptions are, however, neither intrinsic nor crucial to our basic conclusions. Just so long as successively shrinking net demand curves characterize the demands of more distant buyers, our conclusions hold regardless of other facets of buyer distribution and/or tastes.[8] In the

8. What has been shown in the present chapter is that the Case III requirement cannot apply if the monopolist faces innumerably many weaker and weaker markets along with a strongest market, the situation intrinsic to a space economy, *ceteris paribus*. And see F. M.

space economy, price discrimination always yields greater output for the spatial monopolist than does simple f.o.b. pricing.

References

[1] R. Battalio and R. B. Ekelund. "Output Change Under Third Degree Discrimination." *Southern Economic Journal,* 1972, 39:285–90.

[2] M. L. Greenhut and H. Ohta, "Monopoly Output Under Alternative Spatial Pricing Techniques." *American Economic Review,* September 1972, 62:705–13.

[3] A. C. Pigou. *The Economics of Welfare.* 3d edition. London, 1929.

[4] J. Robinson. *The Economics of Imperfect Competition.* London, 1933.

[5] F. M. Scherer. *Industrial Market Structure and Economic Performance.* Chicago, 1970.

Scherer [5, p. 254] where, pursuant to Pigou's nonspatial formulation, he concludes that if demand functions are linear, "output under discrimination will be identical to output under simple uniform-price policy," and where pursuant to Mrs. Robinson's nonspatial analysis, he states "it is impossible to determine whether on balance third degree discrimination increases output."

5. Fundamentals of a theory of spatial monopoly price discrimination

I. *Introduction*

The classical economists asserted that a discriminatory monopoly may not produce more outputs than a simple monopoly. They claimed that the shapes of demand curves alone determine total output under discrimination. In partial contrast, economists who believe the cost of distance is a fundamental variable in microeconomic theory (let us call them the *spatial economists*) contended that a profit-maximizing *spatial* monopolist practices discriminatory or nondiscriminatory pricing on the basis of the shape (not shapes) of the demand curve. Significantly, the references "shape of the demand curve" and "shapes of demand curves" are sharply distinctive. The latter is traditionally used to refer to more than one market, none of which is spatially differentiated, while the former refers to spatial markets. The shape of a demand curve in one market is compared in the classical literature with the shape of a demand curve in another market, as the relative shapes of the demand curves are considered to be of central importance in the theory. On the other hand, the same gross demand curve for all markets and consumers is typically assumed in spatial analyses of discriminatory pricing.[1]

The purpose of this chapter is to complete the transition from classical "spaceless" propositions to the counterpart "spatial" propositions. Assumptions of a constant freight rate per unit of distance and of buyers possessing the same identical gross demand curve are continued. Generally, these buyers are conceived herein as distributed evenly over a plain. What may be called the Hoover-Smithies-Dewey thesis will, accordingly, be synthesized in this chapter. In fact we shall go beyond their thesis and demonstrate that not only may spatial monopolists pursue discriminatory pricing because it is more profitable than nondiscriminatory pricing, but such sellers will always be able to discriminate against nearer buyers *regardless of the shape of the buyer's gross demand curve.*

1. See Hoover [5], Smithies [8], Enke [2], Dewey [1], Greenhut [4]. Note, however, that even if the demand curves are apparently the same, they are *in fact* all different so far as the spatial seller is concerned and thus price discrimination is possible. More important, this assumption of identical demands is methodologically required in order to isolate the pure effect of space.

II. *The spatial model*

Our initial analysis is based on the following set of assumptions.[2]

(1) A monopolist is located at an extreme point of a linear market.
(2) Consumers are evenly distributed along an infinitely long line.
(3) All consumers have identical negatively sloping gross demand curves.
(4) The dollar freight rate per unit of distance per unit of quantity is unity.[3]

These assumptions are fundamentally the same as Hoover's [5], and exactly the same as Smithies [8].[4] They signify that the delivered price p actually paid by a consumer at distance K is greater by the freight costs to that point than the price which the seller receives. The net price for a given unit of sales q is, therefore, given by $p - K$, where K measures distance and, because of assumption (4), also the freight rate per unit sold over that distance. These assumptions indicate that spatial sellers face innumerably many different demand curves, even when buyers have identical gross demand curves. Effective-net demands differ as to intercept values, not as to slope. This condition alone suffices to establish the basis for price discrimination. However, the mere fact that the potential for price discrimination exists does not ipso facto require price discrimination. Among other interests, this chapter inquires into the conditions which make price discrimination most likely to occur. Its basic conclusion will continue to be that price discrimination is as characteristic of and as intrinsic to a spatial economy as is the simple single mill price of the perfectly competitive firms of the spaceless economy.

Hoover's propositions

Hoover assumed marginal production costs (MC) to be constant.[5] He then claimed that profit maximization under spatial discrimination requires equality *in each submarket* of marginal revenue net of freight costs (i.e., net MR) with the marginal cost of production. From the profit-

2. Cf. Hoover [5, p. 183], Smithies [8, p. 63], Enke [2], Greenhut [4], Dewey [1, p. 51], among others.

3. Let $F = F(K,q)$ stand for the freight cost F as a function of distance K and quantity q shipped. Then we mean that $F/Kq = 1$ by assumption (4).

4. Hoover [5, p. 183]; Smithies [8, p. 65]. In contrast, a location model by Hotelling [6] restricted assumption (3) to the case of infinitely inelastic demand, with assumption (1) being replaced by a duopoly market. The Hoover-Smithies model and the Hotelling model are also different in design. The former is concerned with the *pricing policy* of the monopolist, while the latter is concerned with the location of the duopolists. For a concise summary of Hotelling's model, see Ferguson [3, pp. 310–12].

5. Hoover [5, p. 183]. All submarkets are implicitly assumed to be independent of each other as a result of spatial separation as will be evident later in equation (5-1). Cf. Simkin [7, pp. 4–5].

maximizing condition (net MR = MC), Hoover derived the specific relationship between the rate of freight absorption and a change in the elasticity of demand along any given demand curve, viz., the shape of the demand curve. In particular, he set forth certain propositions which may be described simply as follows. Discrimination is practiced against distant buyers when (1) elasticity is constant throughout the individual buyers' demand curves, or (2) the change in elasticity is direct but less than proportional to the change in price, or (3) the elasticity changes oppositely to the change in price. No discrimination occurs when (4) elasticity varies directly and proportionately with price. Finally, discrimination proceeds against nearer buyers whenever (5) the elasticity changes directly and more than proportionately to the change in price.

Let us repeat these propositions in terms which subsequently will facilitate analysis of their meaning and significance:

(1) Freight absorption in the seller's f.o.b. mill price is *negative* when the demand curve is of *constant elasticity* throughout the buyer's demand curve. *Negative freight absorption* means that $1 - dp/dK < 0$, since the expression dp/dK measures the rate at which an increase in the cost of distance changes price.[6]

(2) Freight absorption is also *negative* if *elasticity e decreases less rapidly than price.* Decreasing elasticity means $de/dp > 0$, or alternatively $de/dq < 0$. *Decreasing less rapidly than price* requires $(de/dp)(p/e) < 1$.

(3) Freight absorption is also *negative* when the demand curve is of *increasing elasticity,* where by increasing elasticity we mean $de/dp < 0$ (or alternatively, $de/dq > 0$).[7]

(4) Freight absorption is *nonexistent,* i.e., there is no freight absorption and hence no price discrimination, if *elasticity decreases proportionately with price.* This elasticity requires $(de/dp)(p/e) = 1$.

(5) Freight absorption is *positive,* i.e., $1 - dp/dK > 0$, if elasticity *decreases more rapidly than does price,* i.e., if $(de/dp)(p/e) > 1$.

Let us now cast these propositions in a form that is easily subject to analysis. The framework of thought that follows will reflect Hoover's conclusions and conform perfectly to his explicit assumptions. At the same time, certain errors implicit to his propositions will be uncovered.

6. We may define $1 - dp/dK$ or $(dK - dp)/dK$ as the rate of freight absorption because it measures the rate at which an increase in the cost of distance, i.e., dK, is absorbed by the seller via a change in price, i.e., dp. The negativity of this rate implies that the firm "exacts phantom freight" (Dewey [1, p. 50]) instead of absorbing freight.

7. Increasing (or decreasing) elasticity is thus expressed here in connection with a decrease in price, though either a decrease in price (increase in quantity) or an increase in price (decrease in quantity) can be viewed once the definition is given.

Equation (5-1) is set forth as our point of departure. It applies to any profit-maximizing firm which considers $K \geqslant 0$:[8]

$$p\left(1-\frac{1}{e}\right)-K=c, \tag{5-1}$$

where c is the marginal cost of production, p stands for the delivered price of the seller, and e is defined as $-(dq/dp)(p/q)$. Because $c \geqslant 0$, e is assumed always to be equal to or greater than unity.[9] Taking into consideration the definition of elasticity, differentiating (5-1) with respect to K, applying the chain rule, and effecting appropriate substitutions yield[(I)]

$$\frac{dp}{dK}=\frac{e}{\epsilon-(1-e)}, \quad \text{where } \epsilon=\frac{de}{dp}\frac{p}{e}. \tag{5-1)$'$ (}$$

It follows as we shall see in detail below that

$$\frac{dp}{dK}>1 \quad \text{if and only if} \quad \epsilon=0,\ 0<\epsilon<1, \quad \text{or} \quad 1-e<\epsilon<0. \tag{a}$$

$$\frac{dp}{dK}=1 \quad \text{if and only if} \quad \epsilon=1. \tag{b}$$

$$0<\frac{dp}{dK}<1 \quad \text{if and only if} \quad \epsilon>1. \tag{c}$$

$$\frac{dp}{dK}<0 \quad \text{if and only if} \quad 1-e>\epsilon. \tag{d}$$

Before considering the meaning of these relations, observe that the requirement in (d) is equivalent to specifying $\epsilon<0$, since, as noted above, $e \geqslant 1$. However, because the $\epsilon<0$ in (d) is an even more negative epsilon than its counterpart in (a), where $1-e<\epsilon<0$, we have set forth only the general form $1-e>\epsilon$ in (d) without including the implicit specification $\epsilon<0$.

Relation (a) above proves Hoover's propositions (1) and (2);[10] it also

8. Equation (5-1) relates to a system of equations set forth in Appendix II to this chapter for interested readers.

9. Production cost is a function of the firm's total output Q; hence, marginal cost is independent of the variable K. Cf. Smithies [8, p. 65]. Note also that marginal cost c in (5-1) may be zero, or, of course, positive.

10. If elasticity is constant, $\epsilon=(de/dp)(p/e)=0$; then, $e/[\epsilon-(1-e)]>1$; and hence, $dp/dK>1$. If elasticity decreases less rapidly than price, then $0<\epsilon<1$ and $e/[\epsilon-(1-e)]>1$; again $dp/dK>1$. The third possibility noted in (a), i.e., $(1-e)<\epsilon<0$, will be discussed in notes 11 and 12.

proves Hoover's proposition (3) in a rigorous sense, although Hoover erroneously claimed that $dp/dK > 1$ simply when $\epsilon < 0$.[11] His conclusion was not constrained by the necessary condition which we set forth above, namely, that $(1 - e) < \epsilon < 0$.[12] Hoover's proposition (3), therefore, must be revised by this constraint. In turn, relations (b) and (c) prove his propositions (4) and (5) respectively.[13] The last possibility, case (d), was not recognized by Hoover; however, it may well be dropped because it would involve an economically unstable market.[14]

Hoover's propositions, or the corresponding more precise relations (a), (b), and (c) which we formulated on the basis of his pioneering work, may be combined advantageously with his stipulation that sellers could be expected to discriminate only against buyers located near to (not distant from) the seller.[15] This combination points to propositions (H-1) and (H-2) [or (GO-1) and (H-2)] as a means of summarizing Hoover's (or the modified) theory simply and completely:

(H-1) The spatial monopolist will sell f.o.b. mill if $\epsilon \leqslant 1$.
(GO-1) The spatial monopolist will sell f.o.b. mill if $1 - e < \epsilon \leqslant 1$.
(H-2) The spatial monopolist will absorb freight if $\epsilon > 1$.

Hoover's *original* propositions (1) and (5) are readily summarized by (H-1) and (H-2), whereas relations (a), (b), and (c) are best summarized

11. Observe that $(1 - e)$ might be numerically greater than $\epsilon < 0$. If so, $e/[\epsilon - (1 - e)]$ would be negative and $dp/dK < 0$, not > 0.

12. If $(1 - e) < \epsilon < 0$, then $e/[\epsilon - (1 - e)] > 1$, so $dp/dK > 1$.

13. If $\epsilon = 1$, then $e/[\epsilon - (1 - e)] = 1$; no freight absorption exists in this case. If $\epsilon > 1$, then $0 < e/[\epsilon - (1 - e)] < 1$, and freight absorption is positive.

14. It can be shown that $dp/dK < 0$ if and only if the marginal revenue for each market is increasing. See relation (5-6)′ below. Hence, given marginal cost $c(Q)$ and distance K (which is independent of each market's demand q), the marginal revenue curve cuts the $c(Q) + K$ curve from below when $1 - e > \epsilon$. *The stability condition for a market is not satisfied.* It can also be shown that increasing elasticity, i.e., $\epsilon < 0$, implies $dp/dK < 0$ and hence instability when $c = 0$. To see this, assume $\epsilon < 0$ and $e_0 = 1$, where e_0 stands for elasticity when $K = 0$. (Note that $e_0 = 1$ is an equivalent assumption to $c = 0$, because $e = 1$ if and only if $c = 0$ in deference to the requirement $p(1 - 1/e) = c$.) But also assume the null hypothesis $dp/dK > 0$. Then, as K increases, p must increase. An increase in p implies in turn a decrease in elasticity, since $\epsilon < 0$ by assumption. But this is an impossibility because $e_0 = 1$ and we must have $e \geqslant 1$. In other words, the null hypothesis must be rejected and the contrary to $dp/dK > 0$ applies. Proposition (d) *must,* therefore, apply whenever $c = 0$ and $\epsilon < 0$. The condition shown in relation (a) of $dp/dK > 1$ when $1 - e < \epsilon < 0$ cannot hold in the case of zero marginal cost, since dp/dK must be < 0 and hence we must have $1 - e > \epsilon$, not $1 - e < \epsilon$.

To sum up, if marginal cost is zero and elasticity is increasing, freight absorption would be excessive (i.e., $dp/dK < 0$). Moreover, when $\epsilon < 0$ and MC = 0, MR can never be zero unless it is increasing with increased output. What would appear as an equilibrium position (i.e., MR = MC) under zero variable cost of production conditions is therefore a position of instability, given increasingly elastic demand. But having recorded this alternative possibility, let us henceforth relegate it to the category of the "unexpected-unimportant."

15. Hoover [5, p. 186].

by (GO-1) and (H-2). The difference between (H-1) and (GO-1) should be manifest. Nevertheless, let it be repeated that the former is left unconstrained by the stability requirement $1 - e < \epsilon$, and the latter is corrected to include this constraint. The condition $1 - e < \epsilon \leqslant 1$ in (GO-1) thus involves all cases of increasing elasticity subject to the stability condition ($1 - e < \epsilon < 0$), as well as the cases of constant elasticity ($\epsilon = 0$), slowly decreasing elasticity (where $0 < \epsilon < 1$), and proportionately decreasing elasticity ($\epsilon = 1$). Under any one of these conditions, the firm will only be able to sell f.o.b. mill, since negative freight absorption is ruled out in light of the resale possibility. The condition $\epsilon > 1$ in (H-2) points in turn to the case where elasticity decreases more rapidly than does price, and hence the firm absorbs freight.

Smithies' and Dewey's propositions

Under basically the same set of assumptions as that specified by Hoover, Professor Smithies [8] derived propositions (S-1) and (S-2):

(S-1) The monopolist will sell f.o.b. mill if the logarithmic demand curve is linear or concave.[16]

(S-2) The monopolist will absorb freight if the logarithmic demand curve is convex.

At first glance, Smithies' propositions (S-1) and (S-2) appear to be different from Hoover's propositions (H-1) and (H-2). However, this is not the case. Sets (H-1), (H-2) and (S-1), (S-2) are in fact equivalent, as is shown in Appendix I below. In a similar pattern, Dewey [1] claimed:

(D-1) The firm will exact phantom freight (or sell f.o.b. mill) if marginal revenue declines less rapidly than price at each rate of output (or if the marginal revenue curve and the demand curve are parallel).

(D-2) The firm will absorb freight if the marginal revenue curve declines more rapidly than average revenue.

16. F.o.b. mill price must, itself, depend on the profit-maximizing principle of marginal revenue (MR) equal to marginal cost (MC). To appreciate this, suppose the buyers' demand functions are discontinuous, such as would be the case where the individual domains of each demand are identically defined by $a \leqslant q < \infty$ with $a > 0$; and then suppose $e < 1$. Because MC would be $\geqslant 0$ while MR must be < 0, MR $\neq$ MC. In such case, the monopolist would quote the same delivered price (i.e., the maximum price possible) to each buyer, and he would, accordingly, be discriminating against his nearer buyers in terms of mill price. The conventional marginal principle, based as it is on continuous functions over the relevant nonnegative domain, rules out such a special case even though it is neither theoretically inconceivable nor empirically unobservable. However, as with Hoover (and Smithies), we also rule out $e < 1$, doing so under the paradigm of the marginal analysis and the likely general irrelevance (nonexistence) of $e < 1$ over the entire domain.

The correspondences between these propositions and those of Smithies and Hoover are also shown in Appendix I. For the present, let it suffice to say that the derivative of MR with respect to price is < 1 if and only if $1 - e < \epsilon < 1$, it equals 1 if and only if $\epsilon = 1$, and is > 1 if and only if $\epsilon > 1$. Dewey's framework of thought thus covers essentially the same grounds as did Hoover's and Smithies', even though they looked at the subject from different viewpoints.[17]

There remains the need to go beyond the Hoover, Smithies, and Dewey statements. In this connection, it may at first glance surprise the reader that the relevancy of propositions (H-1), (S-1) – or for that matter (GO-1) – and (D-1) will shortly be disproved *on purely theoretical grounds*. In the process, the thesis that a spatial firm absorbs freight and discriminates against nearer buyers will be reestablished *regardless of the shape of the gross demand curve*.[18] Most important, the denial of (H-1), (S-1), and (D-1) will not be equivalent to contending that the Hoover, Smithies, Dewey propositions are *logically* invalid. Rather, it will serve to support the claim that a basic part of their formulation fails as an explanation of economic relationships in economic space.

III. *The single requirement for price discrimination by a spatial monopolist*

An economically relevant demand function must be subject to the following conditions:

$$f(0) < \infty \quad \text{and} \quad f^{-1}(0) < \infty, \tag{5-2}$$

where f stands for a demand function in terms of price, and f^{-1} is the inverse form of f (provided that f is a one-to-one function). The economic meaning of (5-2) is that the consumer will not consume infinitely large amounts of a good because it is free, nor pay an infinitely high price for economic goods. The latter requirement of (5-2) is sufficient to support the claim above, provided the cost of distance is assumed to be great enough to make freight cost to very distant places by itself equal to the highest price a buyer would pay for the good. To see the significance of (5-2), the constraint $f^{-1}(0) < \infty$ may be considered alone, even though the requirement $f(0) < \infty$ is economically meaningful in itself.

17. We may mention even one more version of the same principle, namely, Stevens and Rydell [9]. They developed the concepts of a demand curve less convex or more convex than a negative exponential. Their interesting development dovetails with Smithies' concepts of a convex or concave logarithmic demand curve.

18. See Hoover [5, p. 186]; also Greenhut [4].

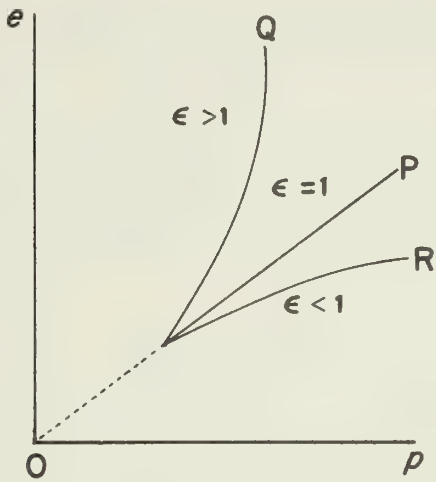


Fig. 5-1 Rates of change in elasticities.

Suppose initially the original basic spatial demand assumption (3) is modified to (3)′:

(3)′ Every consumer has an identical, negatively sloping demand schedule under such constraint that demand vanishes at some critical finite level of price.

Then, to support our thesis, it suffices to show that any demand curve has the constraint specified in (3)′ only if $\epsilon > 1$.[19] As an equivalent demonstration, it will be shown that demand does not vanish for any finite price if $\epsilon \leqslant 1$.

Observe that $\epsilon = 1$ implies a constant e/p, as shown by the P line in Fig. 5-1. It should then be clear that elasticity becomes infinity if and only if price becomes infinity when $\epsilon = 1$. In turn, R in Fig. 5-1 identifies the case where $\epsilon < 1$. This curve therefore indicates that price becomes infinitely high before elasticity approaches infinity when $\epsilon < 1$. But assumption (3)′ implies that elasticity becomes infinity and accordingly demand vanishes while price is still finite. Thus, $\epsilon \leqslant 1$ and assumption (3)′ are incompatible. As the converse, it is obvious that $\epsilon > 1$ implies the relevance of the constraint in assumption (3). It implies this because —as Q in Fig. 5-1 illustrates—elasticity approaches infinity while p is still finite. Subject to the *ceteris paribus* pound, the following fundamental proposition has been established under the assumption that the demand function takes the form of either $\epsilon > 1$, $\epsilon = 1$, or $\epsilon < 1$ *for the whole domain respectively:*

19. If the assumption (3)′ implies $\epsilon > 1$, then via our previous development it also implies that the spatial monopolist will absorb freight.

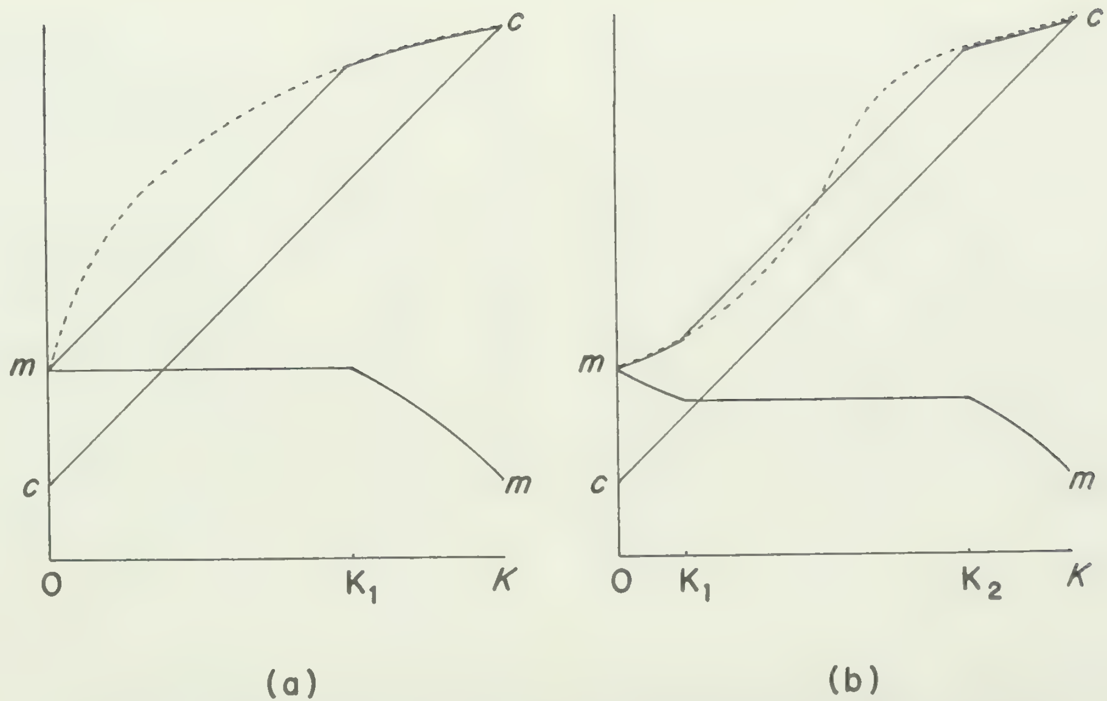


Fig. 5-2 Alternative spatial prices. The monotonically increasing dotted line(s) *mc* stands for the unconstrained optimal delivered prices over space, while the continuous line(s) *mc* stands for the constrained optimal delivered prices. (As noted in the text, since the latter lines are the actually feasible and best ones *given resale possibilities*, they are referred to as constrained but optimal.) The line(s) *cc* stand for marginal cost c plus freight to any applicable distance K. The line(s) *mm* stand for mill prices (= net prices at the seller's mill) applicable to buyers at distance K.

All spatial monopolists selling over a substantial market area always absorb freight if and only if the demand function is constrained as in (3)′.

This proposition suggests that sooner or later—given enough distance—a profit-maximizing spatial monopolist always discriminates in his price against nearer buyers unless constrained otherwise. We shall discuss this rule and elaborate on it in the following final section of this chapter.

IV. *Conclusions*

Conclusion One. Spatial monopolists discriminate generally against nearer buyers regardless of the shape of the demand curve—concave, linear, convex, or some mixture of all. Only the constraint in (3)′ is critically necessary for this result. Indeed, it should be clear that the constraint in (3)′ is intrinsic to any space economy which is sufficiently extensive in length so that natural distance limits would exist to the market area of a spatial monopolist.

Conclusion Two. It has been implicitly assumed that the demand curve is of any type such that $\epsilon > 1$ (or $\epsilon \leqslant 1$) at practically all points, i.e., for practically all p. It is certainly possible for the whole range (or domain) of the assumed demand function to yield the condition $\epsilon > 1$. However, $\epsilon \leqslant 1$ cannot be true for the whole domain given the constraint in (3)′. In particular, elasticity *must* increase more rapidly than price, at least as one approaches the neighborhood of the price intercept, given the constraint in (3)′. This condition alone is sufficient for the spatial monopolist to absorb freight with respect to his most distant buyers even if f.o.b. mill pricing is applied to (some groups of) nearer buyers.

The argument is set forth diagrammatically in Fig. 5-2. In the figure, the dashed *mc* line stands for the delivered price schedule which maximizes profits under the assumption that nearer buyers will not repackage purchased goods and resell them to distant buyers; the continuous line reflects resale possibilities. The dashed line is accordingly referred to as the unconstrained optimal delivered price line; the continuous line is the constrained optimal delivered price line. Fig. 5-2(a) proposes that between O and K_1 the seller prices f.o.b. mill and then discriminates beyond K_1 against the nearer buyers,[20] given the unconstrained optimal

20. To be more rigorous, however, the actual (constrained) delivered price line may start from a point other than point m in Fig. 5-2. Over distances proximate to the seller, the constrained delivered price schedule yields a greater total output than does the unconstrained (optimal) delivered price schedule. Then the *cc* line would shift upward if MC is an increasing function of total output. This rise would, in turn, push the delivered prices

delivered price line *mc*.[21] (A more detailed geometric analysis is presented in Appendix III to this chapter, which explains how the optimal spatial price schedule is derived.)

Another alternative unconstrained price possibility is illustrated in Fig. 5-2(b),[22] which prompts the seller to price f.o.b. mill between K_1 and K_2 and to discriminate elsewhere. Note in this example that the nearest buyers located between O and K_1 are discriminated against besides, in effect, the buyers located between K_1 and K_2 compared with those beyond K_2. More specifically, as distance increases, the mill price first decreases, then stays constant, and finally decreases again until it approaches the delivered cost level *cc*.

Conclusion Three. It is, therefore, a natural phenomenon in economic space for monopolists to discriminate in price against nearer buyers in favor of very distant buyers. Again, this holds even in unusual demand curve cases (e.g., such as those where the elasticity is constant or increasing over a large relevant portion of the aggregate demand curve viewed by the seller), and because spatial sellers tend to be limited by resale possibilities in the way they may discriminate in economic space. So, price discrimination in space is natural except if sellers are otherwise constrained by law, ethics, mores, or their like. It is also likely to be unidirectional under conditions of perfect knowledge among buyers and homogeneous product sales to each buyer.

Conclusion Four. It follows from our derivations that all spatial monopolists obtain greatest possible profits by practicing varying amounts of freight absorption, depending on the buyers' locations.

(constrained as well as unconstrained) upward. (These particular changes in delivered prices are not shown in Fig. 5-2.)

Assuming the *cc* line remains unchanged, the *constrained* optimal price line might nevertheless begin at a point somewhat higher than *m* in Fig. 5-2(a), and would do so if discrimination is to be begun at some nearer distant point than K_1 for which $\epsilon > 1$. To be sure, by setting forth higher prices, the monopolist apparently loses profits on sales to the nearer submarkets. On the other hand, he gains greater profits on sales to more distant buyers since the actual delivered prices would now be both higher and closer to the *unconstrained* prices. The actual delivered prices may stay higher than those depicted by the kinked curve *mc* in Fig. 5-2(a). But the basic conclusion remains unchanged.

21. The optimal delivered prices are readily obtainable by utilizing the profit maximization relation MR = MC and noting that in the space economy MC consists of both the marginal cost of production and transportation. Thus, $p(1 - 1/e) = c + K$, where p stands for the delivered price, e for elasticity of demand which, in turn, is a unique function of p, c for the marginal cost of production, and K for the marginal cost of transportation, assumed here to be equal to distance units. The delivered price p is, therefore, a function of K and c. For simplicity, assume that marginal cost c is constant, in which event p becomes a unique function of K. But, even if c is assumed to be a function of total output, the basic condition that p is a function of K remains unchanged, since c is a function of Q, and Q in turn is determined as soon as the market size K_0 is obtained.

22. Fig. 5-2(b) is drawn taking into consideration the second paragraph of note 20 above.

Conclusion Five. In classical spaceless economics, price discrimination appears as an arbitrary "take advantage of some buyer" policy which only a few monopolists might adopt. In spatial economics, price discrimination does not reflect the same antisocial tinge. Moreover, in classical spaceless economies, arbitrage possibilities are general. But because spatial price discrimination proceeds naturally in favor of the most distant buyers, no arbitrage profit via resale of goods is typically possible, since the delivered prices to these buyers are still greater than those paid by nearer buyers.

Conclusion Six. In order for price discrimination to occur in the classical spaceless monopoly world, the buyer's demand curves must be intrinsically different from that (or those) of others, and even then, for example, output effects depend on relative demand curve relations. In economic space, the gross demand curve of buyers (i.e., the basic tastes and income of spatial buyers) may be identical, and yet the net demand curves viewed by the seller differ; thus price discrimination tends to occur. These relations, and those recorded above, yield the fundamental proposition of this chapter, which may be stated very simply in summary form. Thus spatial monopolists who are rational profit-maximizers *always* discriminate against distant buyers provided they are not constrained by antitrust laws or consumer-public recognition of and objection to this discrimination. Classical spaceless monopoly price discrimination theory, in effect, requires taste and income differences among consumers in order for discrimination to take place.

Appendix I: *Some other propositions related to Hoover's*

As noted in the text of the chapter, Professor Smithies [8] derived propositions (S-1) and (S-2):

(S-1) The monopolist will sell f.o.b. mill if the logarithmic demand curve is linear or concave.[23]

(S-2) The monopolist will absorb freight if the logarithmic demand curve is convex.

At first glance, Smithies' propositions (S-1) and (S-2) appear to be different from Hoover's propositions (H-1) and (H-2). However, this is not the case. Sets (H-1), (H-2) and (S-1), (S-2) are in fact equivalent, as can readily be shown. Continuing the notation used in the text of the chapter, consider our own version of Professor Smithies propositions:

23. Cf. note 16 above.

(GO-2) The logarithmic demand curve is linear (or concave) if and only if $\epsilon = 1$ (or $\epsilon < 1$).

(GO-3) The logarithmic demand curve is convex if and only if $\epsilon > 1$.

To prove (GO-2), let Smithies' logarithmic demand curve be defined as in [8, pp. 63, 64]:

$$\ln q = f(p). \tag{5-3}$$

Then the linearity and concavity of the above function are respectively definable as

$$f'(p) = (1/q)(dq/dp) = -k, \quad \text{where } k \text{ is a positive constant;} \tag{5-4}$$

$$f''(p) = \frac{-(dq/dp)^2}{q^2} + \frac{1}{q}\frac{d^2q}{dp^2} > 0. \tag{5-5}$$

From below[24] (5-4) is found to be equivalent to

$$e = kp, \quad \text{and} \quad \epsilon = \frac{(de/dp)}{(e/p)} = 1. \tag{5-4$'$}$$

In other words, the logarithmic demand curve is linear if and only if $\epsilon = 1$.[25] From definition, elementary manipulation, and substitution of the result derived for d^2q/dp^2 into (5-5),[(II)] we obtain the relation

$$f''(p) > 0 <=> \epsilon < 1. \tag{5-5$'$}$$

Thus, the logarithmic demand curve is concave if and only if $\epsilon < 1$. This completes a proof for (GO-2).

Proposition (GO-3), in turn, is the corollary of (5-5)′. By appropriate substitution in (5-5), we find

$$f''(p) < 0 <=> \epsilon > 1. \tag{5-5$''$}$$

24. Multiplying the right-hand expressions in (5-4) by $-p$ gives $e = kp$. This equality reappears in logarithmic form as $\ln e = \ln k + \ln p$, differentiation of which in turn yields $(1/e)(de/dp) = 1/p$, i.e., $\epsilon = (p/e)(de/dp) = 1$.

25. It is not necessarily required that the slope of the logarithmic demand function be negative as is given here. It could just as well be positive, even when dq/dp is assumed to be negative. It must in fact be positive when the base of the logarithms is less than the natural base e. However, the conclusion will not be affected no matter what base we may use. And see Stevens and Rydell [9, p. 197].

Thus Smithies' propositions (S-1) and (S-2) are respectively equivalent to Hoover's propositions (H-1) and (H-2). Of course, since (H-1) is not obtained directly from the mathematical roots of the model, (S-1) must be limited similarly. Each of these propositions claims the possibility of discriminating against distant buyers (i.e., $dp/dK > 1$) even if $\epsilon < 1 - e$. This is not to say that Smithies failed to provide the stability condition. In fact, he carefully avoided the case where excessive freight absorption occurs. But, at the same time, he did not specify (or provide for) the stability condition *in connection with* the formulation of his proposition. (S-1) would have to be modified to the form (GO-4) in order to conform to (GO-1):

(GO-4) The monopolist will sell f.o.b. mill if the logarithmic demand curve is linear or concave *such that* $1 - e < \epsilon \leqslant 1$.[26]

It is clear that Hoover's propositions (1) through (5) contain more information than (H-1) and (H-2), as the latter simply summarizes part of the former while propositions (1) through (5) do present the mathematical conditions which would allow price discrimination against distant buyers. In turn, they also contain more information than Smithies' propositions (S-1) and (S-2). But notwithstanding the similarities between (H-1) (H-2) and (S-1) (S-2), Hoover's propositions (H-1) and (H-2) are preferable to Smithies' (S-1) and (S-2). Mathematical sophistication via a logarithmic demand function is not necessary. Moreover, a logarithmic demand function presupposes that if the curve is linear or concave, the quantity demanded will never vanish at any finite price, no matter how high it may be. This is not to say that the logarithmic demand function, or any other unrealistic function, is undesirable per se. It is, however, contended that if an economic conclusion is based on an unrealistic assumption, e.g., an unrealistic portion of the assumed demand function, the conclusion itself relates to an unrealistic situation. Before concluding the discussion of this point further, consider Professor Dewey's theory of spatial pricing.

Dewey's propositions

Dewey [1] claimed without offering proof that:

(D-1) The firm will exact phantom freight (or sell f.o.b. mill) if marginal revenue declines less rapidly than price at each rate of output (or if the marginal revenue curve and the demand curve are parallel).

26. The proviso $1 - e < \epsilon$ is derived from the stability condition and not from the concavity of the demand curve. See proposition (d) and note 14 above.

(D-2) The firm will absorb freight if the marginal revenue curve declines more rapidly than average revenue.

To appreciate the scope of (D-1) and (D-2), as well as to evaluate their correspondence to the Hoover-Smithies propositions, one must initially derive the relationship between marginal revenue and price and ϵ. Specifically, marginal revenue is definable as

$$\mathrm{MR} = p\left(1 - \frac{1}{e}\right). \tag{5-6}$$

Via differentiation, we are able to specify

$$\frac{d(\mathrm{MR})}{dp} = \frac{[e - (1 - \epsilon)]}{e} \qquad (5\text{-}6)'$$

And (5-6)′ requires the relations:

$$0 < \frac{d(\mathrm{MR})}{dp} < 1 \quad \text{if and only if} \quad 1 - e < \epsilon < 1; \tag{5-7}$$

$$\frac{d(\mathrm{MR})}{dp} = 1 \quad \text{if and only if} \quad \epsilon = 1; \tag{5-8}$$

$$\frac{d(\mathrm{MR})}{dp} > 1 \quad \text{if and only if} \quad \epsilon = 1. \tag{5-9}$$

The modified Hoover-Smithies propositions may thus be repeated in Dewey's terms as

(HS-1) The firm will sell f.o.b. mill if $0 < d(\mathrm{MR})/dp \leqslant 1$, i.e., if $1 - e < \epsilon \leqslant 1$.

(HS-2) The firm will absorb freight if $d(\mathrm{MR})/dp > 1$, i.e., if $\epsilon > 1$.

Propositions (HS-1) and (HS-2) are not the same as Dewey's (D-1) and (D-2). In the first place, (HS-1) is limited by the possibility of resale while (D-1) is not, since its first part speaks in terms of negative freight absorption. [1, p. 50]. Of course, this difference is simply due to the resale possibility perspective of Hoover and Smithies. There exists no purely theoretical (mathematical) difference in the Hoover-Smithies-Dewey formulations. In the second place, Dewey's propositions involve a terminological vagueness, particularly when he states that marginal revenue declines more (or less) *rapidly* than price; manifestly, such a decrease could imply $[d(\mathrm{MR})/dp]/(p/\mathrm{MR}) > (<) 1$, and not $d(\mathrm{MR})/dp > (<) 1$.[27] We

27. Suppose $d(\mathrm{MR})/dp = 0.5$ at $p = 3$ and $\mathrm{MR} = 1$. Then $[d(\mathrm{MR})/dp](p/\mathrm{MR}) = 1.5$ and hence greater than unity while $d(\mathrm{MR})/dp$ is less than unity by assumption. Thus the two terms are not equivalent at all.

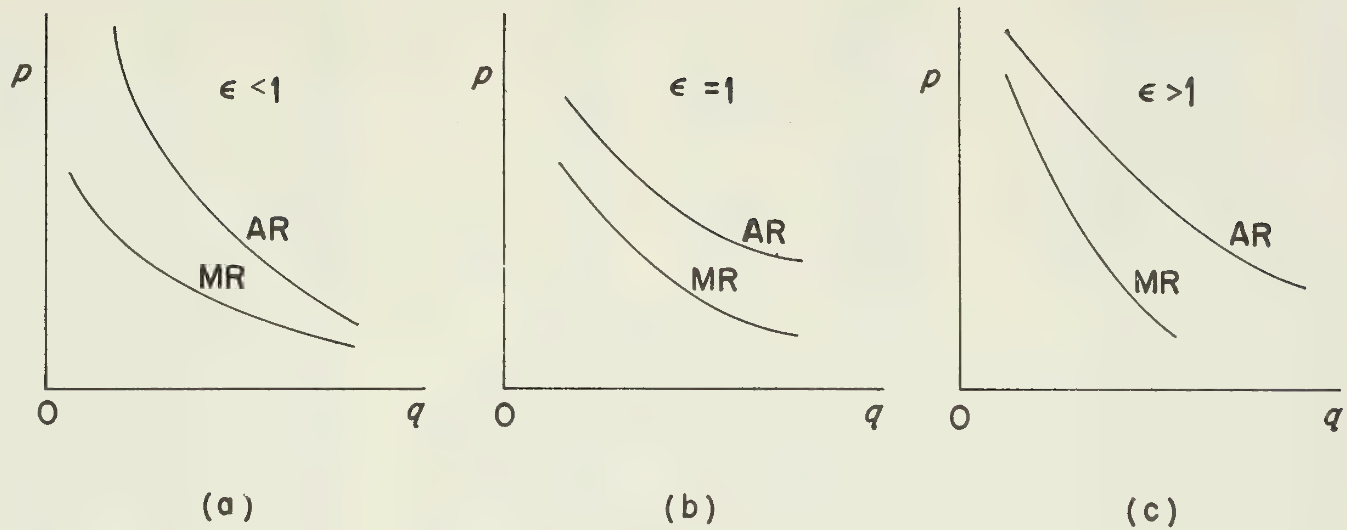


Fig. 5-3 Demand conditions and directions of spatial price discrimination: (a) discrimination against distant buyers; (b) f.o.b.; (c) discrimination against near buyers.

assume he meant the latter, in which event Dewey's propositions are mathematically equivalent to the modified Hoover-Smithies propositions.[28] (Fig. 5-3 illustrates Dewey's propositions.)

Hoover, Smithies, and Dewey therefore proposed substantially equivalent propositions even though each looked at the subject with a different perspective than the others. But the need to go beyond their statements was clear, and it was in this regard that Section III of this chapter was formulated to disprove *on purely theoretical grounds* the general relevancy of propositions (H-1), (S-1)—or for that matter (GO-1) and (GO-4)—and (D-1).

Stevens and Rydell's propositions

Let us briefly review conclusions of Stevens and Rydell on the subject [9]. They arrived at essentially the same propositions as did Hoover, Smithies, and Dewey. Rather distinctively, they compared demand curves in terms of the *negative exponential*. Their positions are these:

(SR-1) The firm will exact phantom freight (or sell f.o.b. mill) if the demand curve is more convex than a negative exponential (or if the demand curve is a negative exponential).

(SR-2) The firm will absorb freight if the demand curve is less convex than a negative exponential.

Here, the demand curve is said to be a negative exponential if $d^2q/dp^2 = (dq/dp)^2/q$ while it is said to be more (less) convex than a negative exponential if $d^2q/dp^2 >(<) (dq/dp)^2/q$.

Now it can easily be shown that

$$\frac{d^2q}{dp^2} \gtrless \frac{(dq/dp)^2}{q} \quad \text{if and only if} \quad f''(p) \gtreqless 0,$$

where $f(p) = \ln q$. Thus it follows that the Stevens and Rydell findings are equivalent to Smithies', and hence equivalent to Hoover's and Dewey's.

Appendix II: *Equation (5-1) and the space economy*

Equation (5-1) in the text of the chapter is based on the following system of equations, where i (= 0, 1, 2, . . . , n) is an integer indicating the distance from the seller to a given buying point.

28. Dewey's propositions are explicitly subject to the stability condition and hence require no modification at all.

$$p_i = f(q_i), \quad \forall i \qquad (5\text{-}10)$$

where p_i is delivered price per unit quantity at the ith consuming point

q_i is quantity of demand at the ith consuming point

$$m_i = p_i - \frac{F_i}{q_i}, \quad \forall i \qquad (5\text{-}11)$$

m_i is mill price per unit quantity applicable for the ith consuming point

F_i is freight cost applicable to the ith consuming point

$$F_i = F_i(q_i, i), \quad \forall i \qquad (5\text{-}12)$$

$$C = g(Q) \qquad (5\text{-}13)$$

C is aggregate cost of production

Q is aggregate (total) quantity produced

$$Q = \sum_{i=0}^{n} q_i \qquad (5\text{-}14)$$

$$R = \sum_{i=0}^{n} m_i q_i \qquad (5\text{-}15)$$

R is aggregate net revenue, net of freight cost

$$\Pi = R - C \qquad (5\text{-}16)$$

Π is profit

$$\frac{\partial \Pi}{\partial q_i} = 0, \quad \forall i \qquad (5\text{-}17)$$

Equation (5-10) stands for a group of identical demand functions, each of which is defined at the ith consuming point. Equation (5-11) provides the mill prices for the different consuming points. Equation (5-12) defines freight cost as a function of the quantity shipped and distance involved. Equation (5-13) is the conventional cost (of production) function. Equation (5-14) states the condition that the total output produced be equal to the aggregate of goods shipped to all relevant consuming points. Equations (5-15) and (5-16) are mere definitional identities, the meaning of which should be evident. Equation (5-17) is a group of first-order conditions for profit maximization. (The system consists of $4(n + 1) + 4$ equations in the same number of unknowns.) If each submarket is independent of any other submarket, the monopolist can

regulate the quantity shipped to *each* so that total profit is maximized. Then, equation (5-17), subject to the rest of the equations, can be rewritten via simple manipulation as

$$\mathrm{MR}_i - \frac{\partial C}{\partial Q} - \frac{\partial F_i}{\partial q_i} = 0, \quad \forall i, \tag{5-17$'$}$$

where MR_i is marginal revenue earned from the ith consuming point. Marginal revenue for any consuming point must, in other words, be equated with marginal cost of production plus marginal cost of transportation applicable to each point in the space. It warrants mention that each consuming point is conceived at this time to have been discretely fixed over the firm's market area. In contrast, our discussion in the text assumed a continuous distribution of buyers for analytical convenience. It should also be noted that (5-1) in the text is, otherwise, equivalent to (5-17)$'$, given the assumption of a unitary freight rate per unit of distance.

Appendix III: *A geometric view of the theory*

The purpose of this appendix is to provide a fundamental method of geometrically deriving the spatial price schedule shown in Fig. 5-2. In doing so, let us first recall equation (5-17)$'$. For simplicity, assume a constant marginal cost of production c and a constant unit freight rate t. Then (5-17)$'$ reappears as

$$\frac{\partial \Pi}{\partial q} = \mathrm{MR} - tK - c = 0, \tag{5-18}$$

which demonstrates that the marginal revenue derived from any market point K must be equated with the marginal costs of producing and shipping the last unit of output to a given point K in the market. It should be stressed that the marginal cost of *transportation* is an increasing function of K, while the marginal cost of *production* is, of course, independent of K.

Marginal revenue MR (as a schedule) is also independent of K. It is, furthermore, redefinable in terms of price p and the elasticity e of demand at any point K in this market. Equation (5-19) therefore applies:

$$\mathrm{MR} = p\left(1 - \frac{1}{e}\right), \quad \forall e \geqslant 1. \tag{5-19}$$

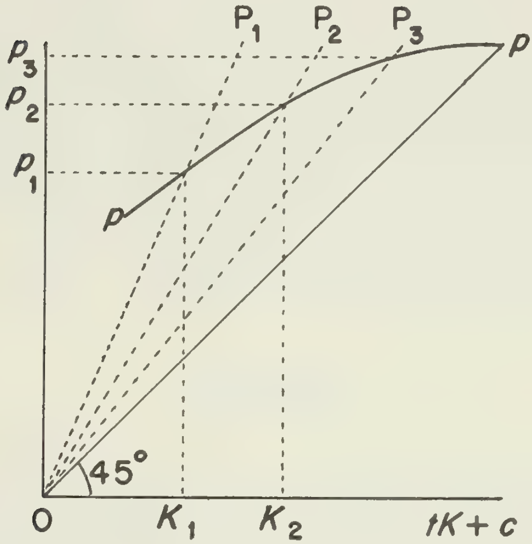


Fig. 5-4 Derivation of optimal spatial prices.

Substituting (5-19) into (5-18) and assuming $c > 0$ yield

$$p = \frac{e}{e-1}(tK + c), \quad \forall e > 1. \tag{5-20}$$

Manifestly, (5-20) can be conceived in Fig. 5-4 as a set of dotted straight lines OP_i with angle greater than 45°. However, elasticity is a single function of price, *given* any specific demand function. Thus, (5-20) is, in fact, a function from p to K. To understand this relationship more fully, consider the following.

Related to a demand function and a specific level of p is a particular value of e. This value e in turn is specifiable via (5-20) as the line OP_1, OP_2, OP_3 of Fig. 5-4. (The subject lines may be called the elasticity identification lines.) The greater the elasticity, the generally further is the elasticity identification line (e.g., OP_3) from the vertical axis. Thus, for example, assume identical demand curves exist at each point in the space whose elasticity increases as price increases. It follows that the higher the price is as a result of distance, the greater is the demand elasticity of the buyer at that distance. In turn, identification of the particular p_i's with the lines OP_i yields the equilibrium points in Fig. 5-4, the locus of which constitutes the optimal price schedule pp over the seller's market space.

The curve pp is normally assumed to be increasing as in Fig. 4-4.

However, a price schedule can be followed in which prices increase more rapidly than the related economic distance. (For example, at distance K_2 in Fig. 5-4, an OP line much steeper than OP_2 may apply, and a concave upward *pp* schedule would characterize the market between distances K_1 and K_2.) But would a monopolist follow such a price schedule simply because the mathematical roots of the demand curves support this schedule? The answer is no, because of the possibility of resale by nearer customers to more distant customers. In such case, therefore, a constrained price schedule, viz., f.o.b. pricing, over space would obtain, as shown in the text, and in detail in [4].

References

[1] D. Dewey. "A Reappraisal of F.O.B. Pricing and Freight Absorption." *Southern Economic Journal,* July 1955, 22:48–54.
[2] S. Enke. "Space and Value." *Quarterly Journal of Economics,* 1942, 56:627–37.
[3] C. E. Ferguson. *Microeconomic Theory.* 2d edition. Homewood, Ill., 1969.
[4] M. L. Greenhut, H. Ohta, and J. Greenhut. "Derivation of Optimal Spatial Prices." *Environment and Planning,* 1974, 6:191–98.
[5] E. M. Hoover. "Spatial Price Discrimination." *Review of Economic Studies,* 1937, 4:182–91.
[6] H. Hotelling. "Stability in Competition." *Economic Journal,* March 1929, 39:41–57.
[7] C. G. F. Simkin. "Some Aspects and Generalization of the Theory of Discrimination." *Review of Economic Studies,* 1948/49, 15:1–13.
[8] A. Smithies. "Monopolistic Price Policy in a Spatial Market." *Econometrica,* 1941, 9:63–73.
[9] B. H. Stevens and C. P. Rydell, "Spatial Demand Theory and Monopoly Price Policy," *Papers, The Regional Science Assn.,* 1966, 17:195–204.

6. Price levels, profits, and outputs under nondiscriminatory and discriminatory spatial pricing

I. *Introduction*

We have traveled a bumpy, winding road from classical spaceless and spatial price discrimination theory to our own version of *what is actually the case.* The bumpiness of the road reflects the indeterminacy of classical spaceless economic theory as to whether output increases, stays the same, or decreases when monopolists discriminate in price. This uncertain underpinning itself reflects the many alternative demand type conceptions admissible in spaceless economic theory. Accordingly, it should not be surprising, we would imagine, for us to undertake analysis of spatial price discrimination along the same lines. In fact, extension of price discrimination theory to include economic space was found to be coterminous with the tendency to pursue unnecessary lines of inquiry, probably a practice which stemmed from the ill-defined set of demand alternatives that had been stressed in the past.

The exhaustive categorization of demand possibilities evident therefore in both classical spaceless and spatial theory had many soporific effects. They found main expression in (two) distinctive, but not well-defined marginal revenue curves. The truth of the matter, as we have seen, is that spatial demand conditions tend always to yield well-defined marginal revenues for discriminating and nondiscriminating monopolists. So output is also well defined, indeed always greater *ceteris paribus* in the case of spatial price discrimination. The implication of the full range of demand possibilities that may be conceived of in a spatial framework does not end at all in the same uncertainty that this range produces in spaceless economic theory.

As soon as one applies economic constraints to spatial demands, it becomes manifest that spatial monopolists *will* discriminate regularly against nearer buyers in favor of the most distant buyers, except of course when the *ceteris paribus* pound is breached by law, custom, and related institutional practices. By adding an *a priori* statement as to empirical

likelihood, spatial price discrimination becomes uni-directional in nature. It is only under exceptional market conditions (as to demand, communications, economic isolation, etc.) that discrimination would proceed in favor of buyers located closest to the seller at the same time that the most distant buyers are also to be favored. The analyses of the past chapters clearly herald the greater profitability of price discrimination over f.o.b. mill pricing and the natural advantage of discrimination being practiced against the most distant buyers.

Certain questions inescapably follow: How do the spatial price structures appear? How much greater are the outputs and the profits of the discriminating firm vis-à-vis the f.o.b. mill seller? And how are these results influenced by the well-defined but alternative demand curve types which our preceding chapters have indicated can (and would) underscore the monopolist's view of the space economy? The standard analysis must be cast aside to answer these questions. And so in the present chapter we shall abandon the beauty and exquisite structure of "general abstract" mathematical models, presenting instead numerical specification of the facets of spatial price discrimination that were just mentioned above. We shall do this by setting forth discriminatory and nondiscriminatory spatial price models and then evaluating them through the demand curve types which previous investigations indicated are relevant to the space economy. Later, in Part III of the text, the effect of free entry on the price practices, outputs, distances, and profits of competitive firms in economic space will be examined.

It is the present assignment, therefore, to reexamine the two spatial pricing systems discussed in the preceding chapters of this book towards the end of determining some of their *specific* salient characteristics. In particular, we shall seek to compare the two spatial delivered-price *schedules* stressing the differences in their values at alternative market points. We shall also specify numerically the difference in the profits and outputs of the two pricing systems. To make the model simple and manageable, but still complete, we shall assume throughout the chapter that (a) buyers are distributed evenly along a line or lines, (b) the freight rate per unit of distance is unity, i.e., the freight cost per unit of output per unit of distance is unitary, (c) the firm's variable production costs are zero,[1] and (d) the buyers' demand functions are identical. Our last assumption means that if it were not for the varying distances which apply with respect to the seller and his buyers, the demand of any buyer would be the same as that of any other buyer as their gross demands are iden-

1. Although this is a simplifying assumption, it is not so unrealistic as would be expected at first glance; and see Cournot [2, p. 60] on this point. Also see Dupuit [3, 4] where he discusses tolls on a bridge as well as railway fares under the same zero marginal cost assumption.

tical. We shall relax assumptions (b), (c) and (d) in the Appendix to this chapter.

II. *The two models*

The first spatial pricing system (model I) consists of the following equation system:

$$p = f(q), \tag{6-1}$$

$$x = K, \tag{6-2}$$

$$\left(\frac{dp}{dq}\right) q + p = x, \tag{6-3}$$

where p = the delivered price, x = the freight rate, q = the demand (at any point on a line), and K = the distance from the seller to any buyer. Equation (6-1) is a demand or demand density function applicable at any point on a line (or lines).[2] Equation (6-2) is the freight cost equation; it signifies that freight rates are proportional to distance.[3] Equation (6-3) involves profit maximization at each point in the market, since marginal revenue is required to be equal to the relevant freight cost (rate). (Recall

2. It must be noted that demand is being measured here somewhat differently than was the practice in Chapter 4. In that chapter, demand was evaluated at each point of a line market and then aggregated. Now demand is being measured only in terms of some interval of a line market, i.e., we now integrate rather than summate. This integration is the reason why (6-1) was referred to as a demand *density* function, and hence why it must be said that at any point in the market *density* can be defined but not *demand.* Though patently different, the two approaches are analogous in another sense, since demand density at a point can be used to represent the unit interval demand. For example, let $q = b - p = b - m - K$ stand for a demand density function at a given point in economic space. Then, the interval demand can be shown as

$$q = \int_K^{K+1} (b - m - K)dK = b - m - K - \tfrac{1}{2}.$$

Thus, since the smallest interval, and hence the smallest integer number is assumed here to be unitary, the number $\frac{1}{2}$ can be disregarded, and the interval *demand* can thus be approximated by the demand *density* function, viz., $q = b - m - K$. It follows that a discrete distribution of buyers requires a point demand function, whereas a continuous distribution of buyers requires an interval demand function. But this interval demand function, we have seen, can be approximated by a demand density function, and hence (6-1) may be referred to as a demand density function.

3. The general freight cost function in form similar to that recorded in Appendix II to Chapter 5 may be represented as $X = g(q, K)$, where X stands for the total cost of shipping output q over distance K. In our model, however, the freight rate per unit of distance, i.e., $(X/q)/K$, is assumed to be unity. Hence, the marginal cost of shipping output is given as the freight rate, i.e., the average freight cost. Both are, therefore, equal to K. I.e., $(X/q)/K = X/qK = 1$ yields $X = qK$, so that $\partial X/\partial q = X/q = K$.

that pursuant to assumption (c) the only variable cost is the freight rate to the different locations of buyers in the market.) Model I thus contains three independent equations in three unknowns: namely, p, q, and x, with distance K treated as a parameter. A determinate price and quantity solution therefore applies to any given point of a line market. By aggregating with respect to K, i.e., $\int pq\, dK$, the dollar value of the aggregate output produced and shipped can be obtained.

The second pricing system (model II) involves equation set (6-4) through (6-7):

$$p = f(q), \tag{6-4}$$

$$p = m + x, \tag{6-5}$$

$$x = K, \tag{6-6}$$

$$\frac{\partial}{\partial m}\left[m \int_0^{K_0} f^{-1}\,(m + K)\; dK\right] = 0. \tag{6-7}$$

The newly introduced unknown m stands for the f.o.b. mill price. The delivered price in this model is defined as the f.o.b. mill price plus the relevant freight rate x. Equation (6-7) is the counterpart to equation (6-3).[4] It assumes that the spatial producer sets his f.o.b. mill price to maximize profits (in this case by maximizing aggregate revenues, since freight costs are paid by the buyer).

The delivered price of the firm is determined under model II by adding to the uniquely derived profit-maximizing f.o.b. mill price of the firm the freight cost applicable to the given point in the market. In the preceding model I, however, no unique profit-maximizing mill price exists, because the mill price depends upon the point in the market area for which the delivered price applies. It will thus be seen that a set of mill prices obtains under model I which may be viewed by subtracting from the delivered prices the freight cost to the location for which the delivered price is derived. When price discrimination applies, buyers purchase commodities for varying mill prices, with the more remotely situated buyers generally benefiting from the practice.[5]

4. Let f^{-1} stand for the inverse function of p. Then, $q = f^{-1}(p) = f^{-1}(m + x) = f^{-1}(m + K)$. Thus, the output q shipped to any distance K is a function of m and K, or more precisely of $m + K$. The profit obtained from shipping q to any distance K is then clearly defined as $mq = mf^{-1}(m + K)$, and the aggregate profit is $\int\, mf^{-1}(m + K)dK$ as reflected in (6-7) above.

5. Cf. Chapter 5 above, where this proposition was established in the backlight of the Hoover-Smithies-Dewey proposition.

III. *Homogeneous demand curves and spatial prices*

Consider the following three specific demand curves:

$$p = \frac{3b}{2} - \sqrt{(2bq)}, \quad \text{a concave demand curve where } \frac{9b}{8} \geqslant q \geqslant 0; \tag{6-8a}$$

$$p = b - q, \quad \text{a linear demand curve where } b \geqslant q \geqslant 0; \tag{6-8b}$$

$$p = \frac{3b}{4} - b^{-1}q^2, \quad \text{a convex demand curve where } \sqrt{3}\left(\frac{b}{2}\right) \geqslant q \geqslant 0. \tag{6-8c}$$

These curves are tangent at the point ($b/2$, $b/2$), the initial condition which led to the numerical forms of (6-8a), (6-8b) and (6-8c).[6] Although the particular numerical specifications of (6-8a) and (6-8c) require tedious computations in determining spatial price, profit, and output schedules (which are relegated to the Mathematical Notes at the end of the text), many advantages result therefrom: (i) It becomes possible to specify the impact of economic space on what (if not for space) would have been an identity of prices and output. (ii) The numerical specifications also help simplify the analysis. Finally (iii) the analysis will be seen to involve evaluation of integrated functions, which functions cannot be resolved without specification of all coefficients.[7]

6. A more general form of the concave demand curve is $p = \beta - \alpha\sqrt{q}$. However, in order that it be equivalent to the simplified linear curve given in (6-8b) in the text, i.e., equivalent in other respects than shape, the same equilibrium point is required for it and the linear curve under the condition of spaceless and costless equilibrium. Specifically, we required its tangency to the linear curve at the point ($b/2$, $b/2$) with the slope $dp/dq = -1$. When these two conditions (repeated below) are satisfied, α and β are uniquely determined: $p = b/2$ at $q = b/2$, and $dp/dq = -1$ at $q = b/2$. Correspondingly, the specific coefficients for the convex case are obtained via the general form $p = \delta - \gamma q^2$.

7. Some integrated functions are evaluated in the subsequent discussion. These functions, in turn, are derived from the marginal revenue (MR), marginal cost (MC) equality, where MC is defined in terms of the distance K. In the case of spatial price discrimination, the equality of MR and MC is given by

$$p\left(1 - \frac{1}{e}\right) = K, \tag{i}$$

where p, the delivered price, is a function of demand quantity, q. If we let f^{-1} be the inverse function from p to q, i.e., $f^{-1}(p) = q$, and then substitute the p given in (i) into the function $f^{-1}(p)$, output q is readily seen to be a function of K. So, the aggregate demand in our model is given by integrating q with respect to sales distance K. This yields:

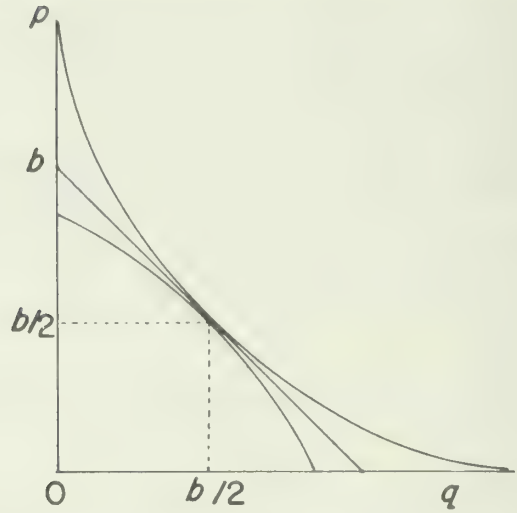


Fig. 6-1 Three alternative demand curves yielding identical nonspatial equilibrium.

It is significant that all three functions conform to the proposition established in Chapter 5; i.e., *all three functions have the property of (monotonically) decreasing elasticity.* Demand curves of constant elasticity (e.g., the rectangular hyperbola) and of increasing elasticity (e.g., a portion of an ellipse) are thus ruled out as is also the decreasing elasticity curve whose change in elasticity is less than proportional to the change in price; theoretical reasons for these exclusions have already been given.[8] Moreover, demand curves of decreasing elasticities are probably the empirically most realistic curves. It is therefore proposed that the three specifications made above are fully general representations of the economically relevant concave, linear, and convex demand curves, so far as the subject matter of this book is concerned.

Profit-maximizing values may be derived for each model in similar ways. Using the concave demand curve for illustrative purposes, and beginning with model I, the profit-maximizing distance K_0^* is easily determined to be $3b/2$, where the asterisk symbol-*-refers (as before) to the profit-maximizing value.(I) In turn, the profit-maximizing revenue R^*,(II) the maximum profit value Π^*,(III) the profit-maximizing output

---

$$\int_0^{K_0} q\,dK = \int_0^{K_0} f^{-1}(p)\,dK = \int_0^{K_0} f^{-1}\left(\frac{eK}{e-1}\right)dK. \qquad \text{(ii)}$$

Obviously the exact value of this function cannot be determined vis-à-vis that which would be derived for the other spatial price model unless the function f^{-1} and hence f is specified.

8. Chapter 5 above, Sec. III.

$Q^{*,(IV)}$ and the profit-maximizing delivered prices $p^{*(V)}$ are readily derived.

In different order for model II, the profit-maximizing mill price (m^*) and distance $(K_0{}^*)$ under concave demand are derived.(VI) Also, the profit-maximizing Q^* and the profit-maximizing R^* (which in this model is the same as the maximum profit Π^*) are specifiable.(VII) Finally, the set of delivered prices over the space, that is, $m^* + K$, is obtained.(VIII)

All results, including those similarly derivable for the other demand curves, are recorded in Table 6-1. Only a preliminary statement on the results recorded in Table 6-1 is set forth immediately below; detailed

Table 6-1 *Spatial prices, distances, profits and outputs*

	Model	m^*	$K_0{}^*\ (= k^*)$	Π^*	Q^*	p^*
Concave demand	I	$\frac{1}{2}b - \frac{1}{3}K$	$\frac{3}{2}b$	$\frac{3}{32}b^3$	$\frac{1}{4}b^2$	$\frac{1}{2}b + \frac{2}{3}K$
	II	$\frac{3}{8}b$	$\frac{9}{8}b$	$\frac{1}{2}\left(\frac{3}{4}\right)^6 b^3$	$\left(\frac{3}{4}\right)^5 b^2$	$\frac{3}{8}b + K$
Linear demand	I	$\frac{1}{2}b - \frac{1}{2}K$	b	$\frac{1}{12}b^3$	$\frac{1}{4}b^2$	$\frac{1}{2}b + \frac{1}{2}K$
	II	$\frac{1}{3}b$	$\frac{2}{3}b$	$\frac{2}{27}b^3$	$\frac{2}{9}b^2$	$\frac{1}{3}b + K$
Convex demand	I	$\frac{1}{2}b - \frac{2}{3}K$	$\frac{3}{4}b$	$\frac{3}{40}b^3$	$\frac{1}{4}b^2$	$\frac{1}{2}b + \frac{1}{3}K$
	II	$\frac{3}{10}b$	$\frac{9}{20}b$	$\left(\frac{3}{10}\right)^3 \sqrt{5}(b^3)$	$\left(\frac{3}{10}\right)^2 \sqrt{5}(b^2)$	$\frac{3}{10}b + K$

discussion of the implications of the table is reserved for the concluding section of the chapter. For the moment, consider only the following two results:

(1) The m^* for model I (which applies to any of the demand curves specified above) is a decreasing function of distance K. In other words, buyers nearer to the seller are required to pay higher mill prices. In this sense, the buyer more proximate to the seller than another buyer is discriminated against under the requirements of model I. On the other hand, the profit-maximizing mill price is, by definition of the f.o.b. price system, uniquely given under model II.

(2) The other vital comparison demanding stress at this time is that the mill price or mill price schedule under model I is mechanically derived from the profit-maximizing delivered prices. The strategic variable for model I is, in other words, the delivered price, while the key variable for model II is the mill price.

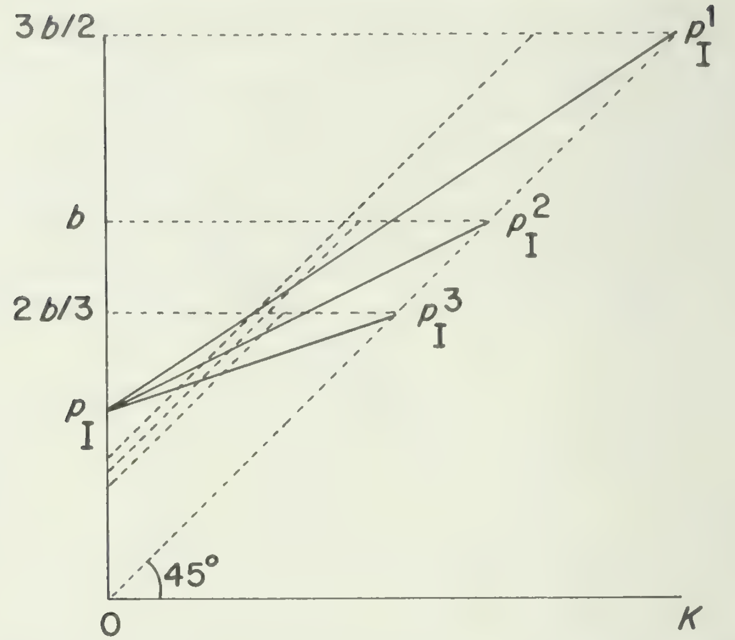


Fig. 6-2 Spatial price schedules under alternative demand conditions and alternative pricing practices. The lines $p_1p_I^1$, $p_1p_I^2$, and $p_1p_I^3$ represent the delivered prices derived from model I under the concave, linear, and convex demand curves, respectively as a function of distance K. Broken lines represent corresponding delivered prices for model II, respectively under concave, linear, and convex demand curves. A similar diagram for the linear demand case and for the circular market case appears in Beckmann [1], pp. 31–32.

IV. *Conclusions on homogeneous demand curves and spatial prices*

Column p^* of Table 6-1 indicates that the more distant buyers are charged lower mill prices under model I; hence discrimination is against the proximate buyers. The same result is shown in Fig. 6-2. Moreover, Table 6-1 and Fig. 6-2 reveal that discriminatory pricing provides a greater (profit-maximizing) sales distance for the seller than does straight f.o.b. mill pricing.

Column K_0^* of Table 6-1 also reveals the dependence of sales distances on demand curve shapes. The linear demand curve, so typically stressed in the literature on spatial markets,[9] is thereby revealed to offer a smaller

9. For example, see [8], [9], and [10].

profit-maximizing distance than the concave demand curve, but a greater trading distance than the convex demand curve.[10]

Column Π^* indicates that the discriminatory pricing system is the more profitable regardless of whether demand is concave, linear, or convex, and thus regardless of the elasticity involved. Profits for the concave, linear, and convex demand alternatives are repeated in a slightly different (rounded) form in Table 6-2 for easier comparisons.

Table 6-2 *Profits respectively shown for models I and II*

CONCAVE DEMAND	LINEAR DEMAND	CONVEX DEMAND
$\frac{3}{32}b^3 > \frac{2.8}{32}b^3$	$\frac{1}{12}b^3 > \frac{2}{27}b^3$	$\frac{3}{40}b^3 > \frac{2.4}{40}b^3$

Column Q^* indicates that output is greater under discriminatory pricing than it is under nondiscriminatory f.o.b. mill pricing. And again, this result applies respectively to each of the demand curves considered in this chapter. Significantly, Q^* is the same for model I under each demand curve. However, as has already been noted, m^* is greater for distances $K > O$, while $K_0{}^*$ is greater and hence Π^* is also greater respectively for concave, linear, and convex demands.

Column p^* indicates that the rate of freight absorption in model I, defined as $1 - dp/dK$, is $\frac{1}{3}$ in the concave case, $\frac{1}{2}$ in the linear case, and $\frac{2}{3}$ in the convex case. These freight absorption *relations* also apply *generally* to concave, linear, and convex demands.[(IX)] It can be shown, in other words, that whenever these demand curves are of decreasing elasticity and are tangent at the spaceless profit-maximizing price, but—unlike Fig. 6-1—are each constrained to the extent that the same maximum-price value (i.e., price limit) applies to each, the absorption rate again is greatest for convex demand and least for concave demand, *ceteris paribus*.[11] In this same connection, note that the rate of freight absorption under model II is nominally zero for all three cases, since actual freight costs are being added to the given profit-maximizing mill price. In de-

10. See [5, pp. 304–8] where, by another derivation, it is shown that for any given distance K, price is greater under concave demand than under linear demand than under convex demand. We find the same result via Fig. 6-2 and Table 6-1, as well as the result that a concave demand curve of the type (6-8a) in this chapter will yield the greatest market area size with the correspondingly highest prices vis-à-vis the curves derived from functions (6-8b) and (6-8c) above.

11. The *ceteris paribus* condition assumes that industrial location would be unaffected by the shape of the demand curve. To understand the scope of this condition, note that if the same maximum price is assumed for each demand curve, convex demand would promote a short-run *localization* of industry and concave demand a *dispersion* of industry [6, chap. 13]. In the long run, however, dispersion of industry tends to arise under both demand curves. In other words, our statement above applies basically without qualification when the long run is assumed. See [5, chap. 6].

riving the mill price, however, a price smaller than the nonspatial price is obtained. The nonspatial price $b/2$ is the greatest, as the spatial prices are $3b/8$, $b/3$, $3b/10$. Some freight is, therefore, absorbed in the process of establishing the spatial price under the given demand situations. Pursuant to the definition given below,[12] the rate of freight absorption under the conditions of model II amounts to $\frac{1}{9}$ for concave demand, $\frac{1}{4}$ for linear demand, and $\frac{4}{9}$ for convex demand. Manifestly, the rates of freight absorption are greater in the case of discriminatory than f.o.b. mill pricing, regardless of the shape of the demand curve.[13]

V. *Summary*

The following propositions were set forth previously: (1) A spatial monopolist will practice the discriminatory price system in which distant buyers are not charged as much freight costs as those actually incurred. (2) From the standpoint of social economy, this spatial pricing system yields more outputs than does the f.o.b. mill price system. (3) The discriminatory pricing system provides greater profits for the spatial monopolist.

12. Alternative definitions are available. One would be to note the difference in the mill price for the spaceless and spatial economies and then to divide this absorption (i.e., difference) value by the largest freight rate facing the seller (in our case the cost of shipping to the most distant point he sells to). Pursuant to this definition, the spatial monopolist absorbs freight at the rate of $\frac{1}{9}$, $\frac{1}{4}$, and $\frac{4}{9}$ respectively under concave, linear, and convex demands. Thus, the values in the text follow this definition. An alternative definition involves the calculation of the average freight rate, i.e., total freight costs divided by total output shipped. Total freight costs are obtainable by the integration $\int_0^{K_0} X\,dK$, and the total output by $\int_0^{K_0} q\,dK$. The rate of freight absorption is then definable as the difference between the spaceless and spatial mill prices divided by the average freight rate, not the freight applicable to the most distant point. This average rate of freight absorption can be shown to be $(\frac{3}{4})^4$ for concave demand, $\frac{3}{4}$ for linear demand, and $\frac{10}{9}$ for convex demand. It may, accordingly, be noted that under convex demand the freight absorption is so great that mill price is reduced by even more than the average freight rate.

13. It was shown years ago – [5, pp. 295–98], and see Chapter 2 above – that under linear demand, zero costs, and an even distribution of buyers along a line, the nondiscriminatory profit-maximizing f.o.b. mill price m^* is $b/2 - K/4$, where K is the freight rate to the most distant buyer, i.e., K_0 in the present book. Because the profit-maximizing limit to the firm's market area is readily seen to be set by $m^* + K = b$, we find by appropriate substitution that this K (or more exactly K_0^*) is $2b/3$, so that m^* in terms of b alone is $b/3$ as in column m^* of Table 6-1. This price and the related rate of freight absorption fall in between those applicable under concave and convex demand. In particular, column m^* (or p^*) of Table 6-1 indicates, for the nondiscriminatory case, the mill price relations $3b/8$, $3b/9$, $3b/10$, respectively for concave, linear, and convex demand. Alternatively viewed, given the same initial conditions – namely, that the same prices and quantities would have applied to each demand curve in the absence of spatial costs – the rate of freight absorption varies with demand curves. As indicated previously, it increases from $\frac{1}{9}$ to $\frac{1}{4}$ to $\frac{4}{9}$ respectively for concave, linear, and convex demand. Under discriminatory pricing characterized by model I and an infinite number of buyers distributed over the plain, all of the freight cost to the most distant effective demander of the seller's product is absorbed by the spatial monopolist.

The analysis of Chapter 6, in particular, has stressed the above relations and many others: (4) The discriminatory pricing system involves a lower schedule of delivered prices for more distant buyers than does the f.o.b. mill system. Hence the market sizes under this price system are the greater. These results obtain because the rates of freight absorption are greater under spatial price discrimination for the reason that profits are maximized at each point in the space rather than over the entire space. (5) The different demand curves yield conformable relations regardless of the price practices followed. Not only do the rates of freight absorption vary from one demand curve to another, but they are at their largest values under conditions of convex demand, next under linear demand, and least under concave demand. (6) In turn, profits, outputs, and price levels (hence sales distances) vary according to the demand curve. And here too, the ordinal relations which apply to the subject demand curves remain in a given sequence (opposite to that of the absorption rates) for each pricing system. Contrary to the result obtained in nonspatial economics where no discrimination occurs when demand functions are identical, spatial monopolists who seek maximum profits *would* pursue discriminatory pricing even if each buyer had the same basic demand. Most significant, the public benefits from spatial price discrimination so far as outputs and price are concerned.[14]

A vital question remains: Will *spatial competitors* follow a single, common mill price, as is the finding among economists who abstract from economic space, or do *spatial competitors* also practice discriminatory pricing over the landscape? Related to this is the question whether the form of the discrimination is the same under spatial competition as under spatial monopoly. Answers to these and related questions will be given in Part III of this book.

Appendix: *Heterogeneous demand curves and spatial prices*[*]

Consider the following. If the cost of shipping a unit of the product 1 unit of distance is t, the transportation cost on each unit sold to a buyer j units away from the seller would be jt, where in the present context of

14. The effect of price discrimination on the efficiency of locational distributions also helps determine the answer to the question of "public costs or benefits" from this pricing technique. Appendix II to Chapter 8 enters into this somewhat tangential subject matter for interested readers. The related subject of sizes and shapes of market areas under the alternative pricing techniques is examined in detail in Chapter 11.

[*] This appendix to Chapter 6 has been drawn from writings by John Greenhut in process of being completed for publication elsewhere. The authors wish to acknowledge their thanks for his permission to recast his material for use herein.

thought the variable K is, therefore, replaced by the symbols jt. Given the proposition that under conditions of costless distance a buyer would purchase quantity q_0 for price p_0, this same buyer would purchase quantity q_0 at the net price $m = p_0 - jt$ when distance is costly and paid for by him. From the spatial seller's viewpoint, any buyer's linear spatial demand curve is therefore definable as $m = (a - jt) - bq$. It follows that if a set of buyers are dispersed evenly along a transport route, the spatial demand of the buyer located at the seller's plant is $p = a - bq$, the buyer at a unit distance away has the demand $m = (a - t) - bq$, and the third buyer's demand is $m = (a - 2t) - bq$; in general, the spatial demand schedule of the $j + 1$st buyer along the transport route is $m = (a - jt) - bq$. But instead of a linear demand, assume more inclusively now that the spaceless demand of the buyer located n units from the seller is a polynomial of the form

$$p = -\alpha q^x + \beta, \tag{6-9}$$

where α and $x > 0$, p and $q > 0$, and the parameters β and x respectively determine the price intercept and general concavity of the demand curve. Then when $x < 1$, the demand curve is concave from above, when $x = 1$, the demand curve is linear, and when $x > 1$, the demand curve is convex from above.[(X)]

Equation (6-10) provides the spatial demand curve of the buyer located j units from the seller:

$$m_j = -\alpha q^x + \beta - jt, \quad j = 1, 2, \ldots n. \tag{6-10}$$

Since the marginal revenue curve associated with this demand is

$$\mathrm{MR} = -\alpha(x + 1)q^x + \beta - jt, \tag{6-11}$$

the profit-maximizing condition $\mathrm{MR} = c$ yields the equilibrium output:

$$q_j = \left[\frac{\beta - jt - c}{\alpha(x + 1)}\right]^{1/x}, \tag{6-12}$$

where c (= MC) is determined in the aggregate and is not necessarily constant.[15] The discriminatory delivered price is then given by

15. The equilibrium MC level c is determined at the intersection of MC with aggregate marginal revenue and hence is a function of the demand parameters α and β and the marginal cost curve. It should be noted that studies of spatial pricing typically assume constant marginal costs. However, such an assumption is unnecessary, since effects of a nonconstant marginal cost curve can be seen to be readily analyzable (and see below, note 17).

$$p_j = (-\alpha q_j^x + \beta - jt) + jt \tag{6-13}$$

$$= \frac{1}{x+1}(x\beta + c) + \frac{1}{x+1} jt.$$

Evaluation of delivered price schedules under different demand conditions is easily effected via (6-13). As a departure point, we shall initially revisit Fig. 6-2 except that a common intercept point will now be assumed, in contrast to Fig. 6-1, rather than a common spaceless equilibrium point.

(a) Delivered price schedules when spaceless demands are identical at each point in the space

The relationship between delivered price and distance is linear when the x and β values are constant over all n. The price at the mill (when $j = 0$) is $[1/(x+1)]\,(x\beta + c)$, the slope of the delivered price schedule is $1/(x+1)$, and the market boundary is limited to $\beta - c$, since when $jt = \beta - c$, $p_n = \beta$, $n = (p - c)/t$.

Fig. 6-3 depicts delivered price schedules when $x = \frac{1}{2}$, 1, and 2 respectively, $c = 2$, and intercept points are assumed identical, of value

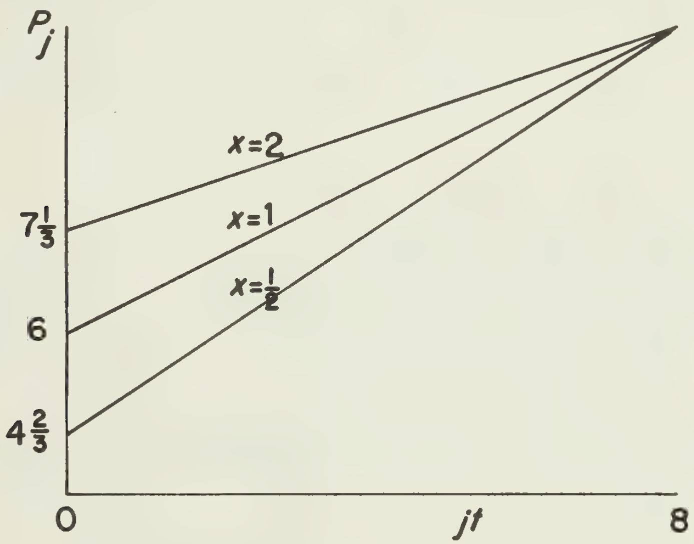


Fig. 6-3 Delivered price schedules from the example: *upper curve,* convex demand; *middle curve,* linear demand; *lower curve,* concave demand.

$\beta = 10$.[16] The more "convex" the demand is (e.g., $x = 2$), the flatter is the slope of the delivered price schedule and the higher is the schedule. Conversely, the more "concave" the demand is (e.g., $x = \frac{1}{2}$), the steeper is the slope and the lower are the prices. It warrants specification, and in fact emphasis, that when x approaches zero, the slope $1/(x + 1)$ approaches unity, i.e, each buyer pays the mill price plus the full cost of shipping the product from the factory to his location. A discriminatory policy may thus resemble nondiscriminatory f.o.b. pricing when x is very low.

Significant weight is attached to the variable x in the light of its impact on the "rate" of discrimination, by which is meant the "amount of freight" that is not charged to (or paid for by) the buyer. Call this particular freight cost the rate of freight absorption and instead of $1 - (dp/dK)$, symbolize it now as

$$f_j = 1 - \frac{p_j - p_0}{jt}, \qquad (6\text{-}14)$$

where p_j is the delivered price j units of distance from the seller, and p_0 is the price charged at the seller's site. When x and β are identical at the mill and distance j, the freight absorption rate reduces to $f_j = x/(x + 1)$. Thus, if $x = 2$, the seller pays $\frac{2}{3}$ of the transportation cost between his location and that of the buyer; an $x = 1$ signifies that the producer pays $\frac{1}{2}$ of the transport cost, etc. Since $df/dx = 1/(x + 1)^2 > 0$, the smaller (larger) x, is the smaller (larger) is f. It follows that the greater the demand "concavity" is, the lower is the freight absorption rate, as is manifest in Fig. 6-1 or 6-2.

(b) Delivered-price schedules when spaceless demands are not identical at each point in the space

The impact on delivered price of the parameter β, i.e., the maximum delivered price a buyer would pay for the good, may be determined from equation (6-13). The greater the value of β is, the higher is the delivered price, *provided the equilibrium value of c is taken as a datum.* For example, assume that the price intercept of the third buyer's spaceless demand curve exceeds that of all other buyers (perhaps because he is wealthier, has a greater need for the product, or is subject to higher prices of substitute goods or lower prices of complementary goods, etc.).

16. The delivered price schedules when $x = \frac{1}{2}$, 1, and 2 are derived from equation (6-11) as:

(i) $p_j = \frac{2}{3}(\frac{1}{2}\beta + c) + \frac{2}{3}jt$ ($x = \frac{1}{2}$, concave demand)
(ii) $p_j = \frac{1}{2}(\beta + c) + \frac{1}{2}jt$ ($x = 1$, linear demand)
(iii) $p_j = \frac{1}{3}(2\beta + c) + \frac{1}{3}jt$ ($x = 2$, convex demand).

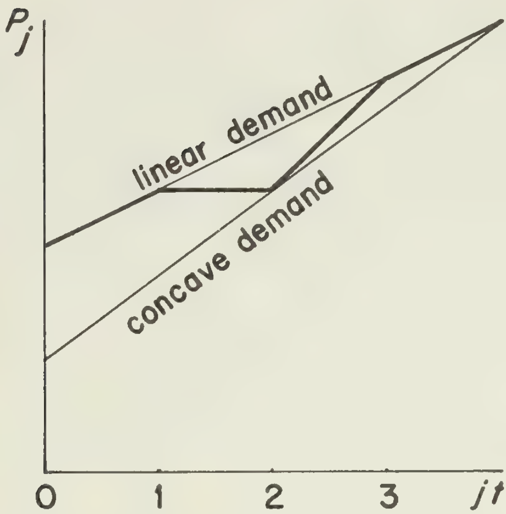


Fig. 6-4 Delivered prices under heterogeneous demands.

Whatever the reason may be, he can be charged a higher delivered price than that which would prevail if the *ceteris paribus* pound applied. It follows that if the spaceless demands are not identical as to price intercept values, the discriminatory delivered-price schedule may not be linear.[17]

The curvature of the delivered-price schedule is influenced by the distance between the seller and the market in which demand for his product is weaker, either in the sense of having a lower intercept value or of being relatively more concave than the demand in adjacent markets. Fig. 6-4 suggests, for example, that the greater the difference is in demand concavities over the space, the greater will be the bumps in the delivered-price schedule. Differences in concavity depend not only on differences in income distribution, but also on the nature of the demand for the product. The more heterogeneous the effective wants of consumers at a given point in space, the greater is that market concavity of demand vis-à-vis other markets, and the bumpier the delivered-price schedule.

17. *The slope coefficient α of the demand curve, on the other hand, does not affect the curvature of the delivered price schedule.* Demand curve slope affects discriminatory spatial pricing only through its impact on the equilibrium marginal cost level c. This result is due to the fact that c is determined at the intersection of marginal cost and aggregate marginal revenue, and since the aggregate marginal revenue is a function of the slopes of the individual demand curves, the equilibrium c is indirectly affected also. Equilibrium MC, in other words, is identical for all locations in the market area. Only the general price level of the schedule may be affected by slopes, owing to the dependence of c on demand curve slopes when marginal costs change with output. The consequence is that the *delivered-price schedule remains linear, ceteris paribus,* regardless of variations in slopes of demand curves at different buying points in the space.

While heterogeneous demands over economic space normally yield nonlinear delivered-price schedules, it was shown [7] that the various demands may combine in such manner as to produce a linear schedule.[18] Because traditional spatial price discrimination theory is based on the assumption of identical spaceless demands over the space, uniform pricing cannot be explained when distances (freight costs) are significant. Yet our theory indicates that uniform pricing could occur under conditions of costly distance. It is unlikely, of course, that varying demands over economic space will "combine" so exactly as to require perfectly uniform delivered-price schedules. However, since it is advantageous for many firms to sell according to simple price policy (e.g., for the firm to absorb a constant part of the freight cost to each buyer), it may be expected that even though varying demands do not "combine" so uniquely as to produce uniform price schedules over space, their relations may yield such close approximation to a simple pattern that it is uneconomical for the firm to administer a complex delivered-price policy. Linear delivered-price schedules may therefore occur with some degree of regularity in free enterprise systems, even when demands vary over economic space. In fact, we shall see in Chapter 8 that the distribution of competitors over economic space also produces nonintuitively obvious results of the order given above.

References

[1] M. Beckmann. *Location Theory.* New York, 1968.

[2] A. A. Cournot. *Researches Into the Mathematical Principles of the Theory of Wealth.* Translated by N. T. Bacon. 1897.

[3] J. Dupuit. "On the Measurement of the Utility of Public Works." *International Economic Papers* (London, 1952), pp. 83–110.

[4] J. Dupuit. "On Tolls and Transport Charges." *International Economic Papers* (London, 1962), pp. 7–31.

[5] M. L. Greenhut. *Plant Location in Theory and in Practice.* Chapel Hill, N.C., 1956.

[6] M. L. Greenhut. *A Theory of the Firm in Economic Space,* 2d printing. Austin, 1974.

[7] J. Greenhut, M. Hwang, and H. Ohta. "Spatial Price Discrimination – Theory, Assumptions, and a Case Example." *Annals of Regional Science,* 1974, 8:70–81.

[8] A. Lösch. *Die räumliche Ordnung der Wirtschaft.* (Jena, 1944).

[9] E. S. Mills and M. R. Lav. "A Model of Market Areas with Free Entry." *Journal of Political Economy,* 1964, 72:278–88.

[10] A. Smithies. "Optimum Location in Spatial Competition." *Journal of Political Economy,* 1941, 49:423–39.

18. Assume four evenly dispersed buyers (A to D) respectively located at the seller's location . . . up to three distance units away. Let c and t be 3 and 1 respectively. Given the demand curves in the left-hand equation set below, the profit-maximizing delivered prices are given in the right-hand set.

A: $p_0 = -2q^{1/2} + 9.$	A: $p_0 = \frac{2}{3}(\frac{1}{2} \cdot 9 + 3) + \frac{2}{3}(0)(1) = 5.$
B: $p_1 = -3q + 6.$	B: $p_1 = \frac{1}{2}(6 + 3) + \frac{1}{2}(1)(1) = 5.$
C: $p_2 = -q^2 + 5.$	C: $p_2 = \frac{1}{3}(2 \cdot 5 + 3) + \frac{1}{3}(2)(1) = 5.$
D: $p_3 = -4q + 4.$	D: $p_3 = \frac{1}{2}(4 + 3) + \frac{1}{2}(3)(1) = 5.$

Part 3. Nondiscriminatory and Discriminatory Pricing The Competitive Oligopoly Case

7. Nondiscriminatory oligopoly pricing: nonspatial compared with f.o.b. mill pricing

I. *Introduction*

The basic principles of simple monopoly pricing and its spatial counterpart f.o.b. mill pricing were set forth in Chapter 2. But analysis of monopoly pricing alone is insufficient to establish adequate answers to the questions *which* sellers discriminate in price and under *what* conditions in the space economy. The present chapter therefore goes beyond comparisons of the simple spaceless and spatial monopoly. It evaluates the impact of competition on spatial pricing, and in the process compares this pricing practice with classical findings of competitive pricing. Because it will later be shown that competition in economic space culminates in an oligopoly market, prices in an oligopoly market are alone considered herein. This chapter accordingly first sets forth selected fundamentals of spaceless nondiscriminatory oligopoly pricing theory, and then examines the spatial form of this pricing system. Chapter 8 concludes the present phase of our study by broadening the view of spatially competitive prices to include analysis and determination of its discriminatory properties.

II. *A classical view of competitive impacts on supply and price*

Cournot's model: sellers and buyers at a point

Assume zero costs and hypothesize as did Cournot [2, pp. 79–80] that sellers regard any rival's supply as fixed. Also assume a linear demand function of the form $p = b - aQ$.[1] Seller One, the first entrant in the

1. Cournot did not assume a specific demand such as linear demand. In effect, basing the subsequent discussion on linear demand reflects Chamberlin's version of Cournot more than Cournot himself. Chamberlin's version is followed simply because it illustrates Cournot's fundamental thesis most readily. But, again, methodological simplicity is alone the reason for emphasizing linear demand here. Cf. Cournot [2, chap. 3] and Chamberlin [1, pp. 32–34].

market, thus maximizes profits by selling one-half of the total that could be sold at the marginal cost (zero) price. A second entrant, Seller Two, would visualize accordingly a potential demand for his product one-half as great in ordinate and abscissa intercept values as that which in total prevails. So the best initial output adjustment for Seller Two is to offer one-half of the greatest amount he would expect to be able to sell if the zero price prevailed. Such a second seller, in other words, supplies one-fourth of the greatest quantity of the product buyers would purchase. But this reduces by one-fourth the abscissa and ordinate extension visualized by Seller One *as his share* of the market. His offerings are cut in turn by half of this one-fourth amount.

Another linkage exists! Contraction by Seller One enlarges the sales potential viewed by Seller Two to the extent of one-eighth of the total quantity that would be sold in the market at the zero price. Corresponding actions then follow. They combine to form the geometric series indicated in Table 7-1.

Table 7-1

	SELLER ONE	SELLER TWO
Initial offering	$\frac{1}{2}$	$\frac{1}{4}$
Change in offering	$-\frac{1}{8}$	$+\frac{1}{16}$
Next change in offering	$-\frac{1}{32}$	$+\frac{1}{64}$
Next change in offering	$-\frac{1}{128}$	$+\frac{1}{256}$
. . .	. . .	. . .

The full geometric series $(a + ar + ar^2 + \cdots + ar^{n-1})$ exists, with the numbers recorded in (7-1)′ and (7-1)″ providing some of the changes in supply that occur with the entry of the second firm. More specifically, (7-1)′ shows supply actions of the second firm, and (7-1)″ resulting changes in supply by the first firm:

$$\frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \cdots \tag{7-1$'$}$$

$$-\frac{1}{8} - \frac{1}{32} - \frac{1}{128} - \cdots \tag{7-1$''$}$$

Each series sums to $[a/(1 - r)] - [ar^n/(1 - r)]$, where $a = \frac{1}{4}$ for Seller Two, $a = -\frac{1}{8}$ for Seller One, and $r = \frac{1}{4}$ for both. Because the second term

$ar^n/(1-r)$ approaches zero as n increases in either case, and the first term $a/(1-r)$ totals up to $\frac{1}{3}$ for (7-1)′ and $-\frac{1}{6}$ for (7-2)″, the second firm sells $\frac{1}{3}$ and the first firm $\frac{1}{2}-\frac{1}{6}=\frac{1}{3}$ of the total demand at the zero price. That is to say, the two firms combine to sell $\frac{2}{3}$ of the demand that would prevail at the zero price. Price therefore falls from $b/2$ to $b/3$ when a competitive duopoly market replaces a monopoly.

Suppose a third firm, Seller Three, enters the market. The original two sellers were supplying $\frac{2}{3}$ of the total market demand that would obtain at the zero price. So Seller Three will offer $\frac{1}{2}$ of the unsatisfied portion, i.e., $\frac{1}{6}$ of the total. But then the other two sellers visualize a shrinkage of demand and reduce their offerings by $\frac{2}{3}$ of $\frac{1}{6}$ or $-\frac{1}{9}$. The altered supply of the original sellers induces Seller Three to increase its offerings by $\frac{1}{18}$, a change in supply that causes a contraction of $\frac{1}{27}$ by the original sellers. The series of changes for Sellers One and Two appears in (7-2)″ and for Seller Three in (7-2)′:

$$\frac{1}{6}+\frac{1}{18}+\frac{1}{54}+\cdots, \qquad (7\text{-}2)'$$

i.e., $a=\frac{1}{6}$, $r=\frac{1}{3}$, so $a/(1-r)=\frac{1}{4}$.

$$-\frac{1}{9}-\frac{1}{27}-\frac{1}{81}-\cdots, \qquad (7\text{-}2)''$$

where $a=-\frac{1}{9}$, $r=\frac{1}{3}$ and $a/(1-r)=-\frac{1}{6}$ from the original $\frac{4}{6}$. Sellers One and Two supply $\frac{3}{6}$ of the total demand at the zero price after a Seller Three has entered the market. In other words, each of three sellers supplies $\frac{1}{4}$ of the demand that would be satisfied by pure competitors, and this sums to $\frac{3}{4}$ of the total market demand. Price descends to $b/4$. It can be shown that four sellers would supply $\frac{4}{5}$ of competitive output at price $b/5$, five sellers would offer $\frac{5}{6}$ of competitive output at price $b/6$, *ad infinitum*. Assuming each seller behaves competitively and entry is open, an oligopoly approaches the competitive output and price as the number of sellers increases without limit.

Introducing positive marginal costs to Cournot's model

When positive marginal costs are assumed to exist as in Cournot [2, p. 85], and demand and other conditions are as defined above, the monopolist's profit-maximizing price is $\frac{1}{2}$ of the ordinate intercept value plus $\frac{1}{2}$ of the marginal cost at the profit-maximizing output. This statement is equivalent to the assertion that price equals $\frac{1}{2}$ the vertical height of the demand curve above marginal costs plus the marginal cost [7,

pp. 30, 32, 55]. Similar relations apply to duopoly and larger numbers of firms.

From the above rules, it follows under the assumption of seller homogeneity and four sellers in the market that any one of the sellers will supply $\frac{1}{4}$ of $\frac{4}{5}$ of the relevant market demand. A geometrical representation of the individual seller's demand curve requires that it be drawn parallel to the market demand curve. The intercept values of the individual firm's demand curve in such four-seller market must, therefore, be $\frac{2}{5}$ that of the market demand curve. Such drawing (following the above rules for a negatively sloped linear demand curve) yields a price in the present example which is greater than the marginal cost value by $\frac{1}{5}$ of the difference between the ordinate intercept and the marginal costs. Also, it provides sales for each seller which are $\frac{1}{4}$ of $\frac{4}{5}$ of that which could be supplied at that marginal cost price.[2]

Another view of oligopolistic interdependencies and pricing

Let r be a representative oligopolistic firm, and assume there exist, in addition to this firm, f homogeneous other firms. The products of the f other firms are slightly differentiated from r's product. Designate r's demand function by (7-3)′, while the demand function for the homogeneous f other firms is given by (7-3)″:

$$p_r = a_0 - b_0 q_r - c_0 \textstyle\sum_{s=1}^{f} q_s; \tag{7-3$'$}$$

$$p_f = a_1 - b_1 q_r - c_1 \textstyle\sum_{s=1}^{f} q_s. \tag{7-3$''$}$$

Though the products of the f firms differ slightly from r, the firms are assumed to be in competition *à la* Cournot. Reaction functions can accordingly be derived. Converting (7-3)′ and (7-3)″ into profit functions, taking the partials in terms of q_r and q_s respectively, and then setting them to zero yields

$$q_r = \frac{a_0}{2b_0} - \frac{c_0}{2b_0} \sum_{s=1}^{f} q_s; \tag{7-4$'$}$$

2. When marginal costs are positively sloped instead of constant at output points in the neighborhood of the MR/MC intersection, a distinction must be drawn between the MR's of the monopolist and the MR's of competitive sellers. For example, if three sellers are in the market, the MC intersection with the allocable MR curve of a particular seller must be viewed. The entry of a fourth seller produces the same results as those applicable to constant marginal costs *except that price would descend a little faster* when MC is rising. Moreover, what is $\frac{1}{5}$ of the relevant demand at the new relevant level of marginal costs is greater than that which would prevail under the assumption of a constant level of marginal costs equal to that which would apply to three-sellers competition in the market, as described initially above.

$$\sum_{s=1}^{f} q_s = \frac{a_1}{2c_1} - \frac{b_1}{2c_1} q_r \quad \text{or its inverse} \quad q_r = \frac{a_1}{b_1} - \frac{2c_1}{b_1} \sum_{s=1}^{f} q_s. \quad (7\text{-}4)''$$

The reaction functions of r and the f firms are depicted in Figs. 7-1 and 7-2. In order to assume a stability solution, the horizontal intercept of one reaction function must be greater than the vertical intercept of the other, viz., $(a_1/b_1) > (a_0/2b_0)$ and $(a_0/c_0) > (a_1/2c_1)$. Given these conditions, their intersection yields in Fig. 7-1 the profit-maximizing (Cournot) supplies for firm r and firms f, and in turn through (7-3)′ and (7-3)″ the profit-maximizing prices.

An alternative model of these interdependencies may be constructed by substituting the value

$$\sum_{s=1}^{f} q_s = \frac{a_1}{c_1} - \frac{b_1}{c_1} q_r - \frac{p_f}{c_1}$$

in (7-3)′, and

$$q_r = \frac{a_0}{b_0} - \frac{c_0}{b_0} \sum_{s=1}^{f} q_s - \frac{p_r}{b_0}$$

in (7-3)″, and evaluating the new forms of (7-3)′ and (7-3)″ under competition. This procedure reflects Edgeworth's approach, in particular the

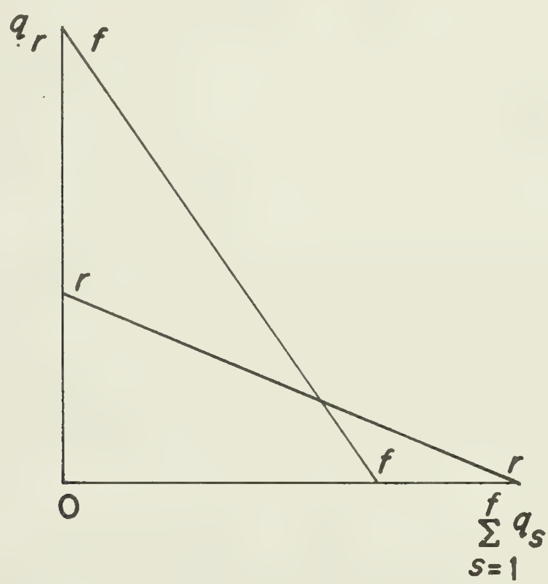


Fig. 7-1 Reaction functions *à la* Cournot.

assumption that the rival's supply *price* is given. The new forms of revenue functions and the corresponding reaction functions are, respectively

$$p_r q_r = \frac{p_r}{c_1 b_0 - c_0 b_1}(a_0 c_1 - a_1 c_0 + c_0 p_f - c_1 p_r), \tag{7-5)$'$(}$$

$$p_f \textstyle\sum_{s=1}^{f} q_s = \frac{p_f}{c_1 b_0 - c_0 b_1}(a_1 b_0 - a_0 b_1 + b_1 p_r - b_0 p_f); \tag{7-5)$''$(}$$

and

$$p_r = \frac{a_0 c_1 - a_1 c_0}{2c_1} + \frac{c_0}{2c_1} p_f, \tag{7-6)$'$(}$$

$$p_r = \frac{a_0 b_1 - a_1 b_0}{b_1} + \frac{2b_0}{b_1} p_f. \tag{7-6)$''$(}$$

The new reaction functions (7-6)′ and (7-6)″ are obtained by taking partials of (7-5)′ and (7-5)″ with respect to p_r and p_f respectively, and setting them to zero. As is clear from Fig. 7-2, it suffices for a stable equilibrium solution in the first quadrant that $(a_0c_1 - a_1c_0)/2c_1 \geqslant (a_0b_1 - a_1b_0)/b_1$ and $2b_0/b_1 > c_0/2c_1$. By elementary operations, these relations are respectively revealed to involve $2a_1b_0c_1 - a_0b_1c_1 - a_1b_1c_0 \geqslant 0$ and $4b_0c_1 - b_1c_0 > 0$, in turn referred to hereinafter as β and α. The stability conditions with nonnegative solutions are then specifiable as $\beta \geqslant 0$, $\alpha > 0$, and $a_0b_1 - a_1b_0 + 2b_0\ \beta/\alpha \geqslant 0$.[(1)] If these conditions are assumed typically (or normally) to be satisfied, the alternative model will also yield a stable equilibrium, as in Fig. 7-2. This stability results notwithstanding the present model's resemblance to the Edgeworth model of competition.

It may be worthwhile to note that (7-5)′ or (7-5)″ specifies p_r and p_f respectively as a simple function of q_r or q_f, since p_f and p_r are assumed by the firm to be fixed. The gist of the matter is that the profit-maximizing equilibrium $\Sigma_{s=1}^{f} q_s$ which is obtained via equation (7-4)′ and (7-4)″ yields a particular average revenue schedule for r in terms of q_r (see the very dark line in Fig. 7-3). Other average revenue schedules for firm r in terms of q_r, resulting from different $\Sigma_{s=1}^{f} q_s$, are also given in Fig. 7-3. As $\Sigma_{s=1}^{f} q_s$ is increased (decreased) by rivals, the p_rq_r schedule shifts downward (upward).

The alternative set of AR schedules are drawn as broken lines in Fig. 7-3; to repeat, these are derived from fixed p_f's. The broken dark-lined p_rq_r curve relates to p_f being at its profit-maximizing value as derived from the intersection of the reaction functions (7-6)′ and (7-6)″. Fig. 7-3

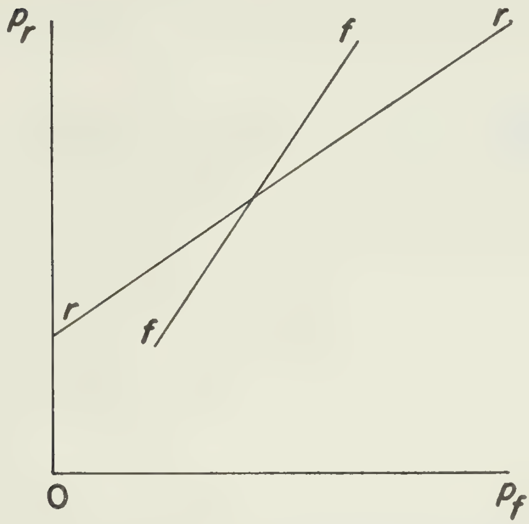


Fig. 7-2 Reaction functions *á la* Edgeworth.

shows that the AR curves with fixed supply by rivals are steeper than those reflective of a fixed price by the rivals. Moreover, it can be shown that the equilibrium price resulting under conditions of competition *à la* Edgeworth is lower than the equilibrium price derived under Cour-

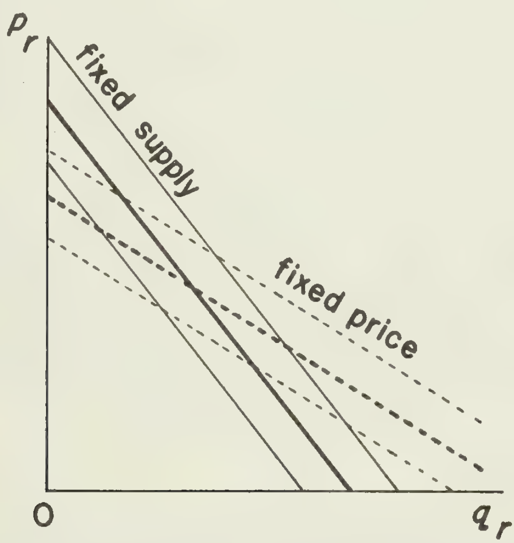


Fig. 7-3 Revenue curves given either fixed supply or fixed price by rivals: *solid lines,* fixed supply Σq_s; *broken lines,* fixed price p_f.

not's assumption. The price intercept of the heavy AR line must accordingly be placed higher than the heavy *broken* line, *ceteris paribus* as in Fig. 7-3.

A third alternative AR curve or curves can be conceived. Suppose firm r believes the f firms will maintain their price with respect to an increase in its price p_r, but would lower their prices below the competitive (Cournot or Edgeworth) profit-maximizing price in order to maintain their proportion of sales. A kinky average revenue curve would then exist.[3] If next another firm enters the market, so that now there are g competitors, where $g = f + 1$, the coefficients a_1/b_1 and c_1 in (7-4)″ decrease in value. The reaction function for these firms as shown in Figs. 7-1 and 7-2 would shift leftward and flatten accordingly at the same time that $\Sigma_{s=1}^{g} q_s$ increases; p_g, p_r, and q_r all would fall. Adjustments to the kinky demand curve will also take place as a result of entry, with change in price eventuating only as described later.

The same relations prevail when competitive entry occurs at a different point in economic space, as where the rival firm r is assumed to be located at a distance from a production center. But this spatial framework brings us to the inquiry of Section III of the chapter. Besides yielding comparable insights through analogy with the spaceless model, the spatial oligopoly framework focuses attention on an entirely new set of forces. In the section that follows, competitive entry at the same location as the rival firm(s) will first be analyzed, and then the impacts of competitive entry at a distance will be determined.

III. *Competitive impacts on spatial supply and price*

Introducing economic space into Cournot's model – buyers distributed over the space: the monopoly case reviewed

For present purposes, consider again the "line" situation where n spatially separated markets of the same size are identifiable, and where the gross demand curve at each market point is linear of the form $p = b - naq$. Let the economic distance or freight rate to any given site on the line market be defined by K. Then, the net linear demand curves are in turn definable by rewriting $p = b - naq$ in terms of mill price and freight cost as

$$q = \frac{(b - m - K)}{na}. \tag{7-7}$$

3. In the conjectured event of an increase in p_r, simply substitute (7-3)″ in (7-3)′ assuming p_f is fixed. In the conjectured event of a decrease in p_r, simply assume either Σq_s is given in (7-3)′ (or the initial – competitive profit-maximizing – proportion of $\Sigma q_s/q_r$ is given in (7-3)′). A kinked $p_r q_r$ curve obtains.

Consequently, the aggregate net demand in the limit (as $n \to \infty$) will be

$$Q = \sum q = \frac{1}{b}\int_0^{K_0}\left(\frac{b-m-K}{a}\right)dK = \frac{\left(bK_0 - mK_0 - \frac{K_0^2}{2}\right)}{ba}. \qquad (7\text{-}8)$$

Assuming zero cost of production, the first-order conditions for profit maximization will yield

$$\Pi = mQ. \qquad (7\text{-}9)$$

$$\frac{\partial \Pi}{\partial m} = 0; \quad m = \frac{b}{2} - \frac{K_0}{4}, \quad \text{and} \qquad (7\text{-}10)$$

$$\frac{\partial \Pi}{\partial K_0} = 0; \quad K_0 = b - m.$$

As derived originally in [4] for the line case under conditions of linear demand, the effect of distance is to force the nondiscriminatory mill price downward by one-fourth of the freight cost to the most distant buying point in the monopolist's market area, or equivalently by one-half the freight cost to the buyer located at the halfway point of the monopolist's market space. Correspondingly, as derived originally in [4] for the plain case, the mill price is $b/3$ and K_0 is $2b/3$.

Another way of describing the cost of distance is to propose that the aggregate effective demand curve in the spaceless world is greater than the aggregate effective spatial demand to the extent of the cost of distance. Recalling further in this connection the delivered price formula $p = b - naq$, and the related intercept points of individual net demands in space, $m_i = b - (i-1)b/n$, where $i = 1, 2, \ldots, n$, it is clear as in Chapter 4 that the sum of all net demand curves yields

$$Q = \frac{i[(b-m) - (i-1)b/2n]}{na}, \quad \forall m \ni m_i > m \geqslant m_{i+1}. \qquad (7\text{-}11)$$

In the costless distance world $n = 1$, the formula reduces to $Q = (b-m)/a$. In the spatial world, the quadratic form appears:

$$\lim_{n\to\infty} Q = \frac{(b-m)^2}{2ab}. \qquad (7\text{-}12)$$

So, the aggregate effective demand in the spatial world is of smaller magnitude than its counterpart in the spaceless world, *ceteris paribus.*

Competitive entry at an existing production center

The impact of entry at the production center in the spatial world corresponds to that in the spaceless world. In particular, reaction functions apply which in reflection of the quadratic aggregate demand curve are themselves quadratic.[II] Because the results are identically counterpart to those described previously in connection with linear demand, further statements are unnecessary. Attention may thus be directed to the more interesting possibility of competitive entry at a distant location.

Competitive entry at a distance

The impact of competitive entry at a distance is sharply distinguishable from classical impacts. Indeed, it is more sharply distinctive than the impact of competitive entry at *the* given production center compared to competitive entry in the classical system of thought, since the term competition in economic space means that a distant firm gains full control over buyers located nearer to it than its rivals. The reason for this impact will be discussed shortly, though details are left to [4, chap. 2, sec. 3].

To view the situation thoroughly, assume for simplicity once again that marginal costs of production are zero, the freight rate is constant, buyers are evenly distributed along an endless line, and the individual negatively sloped demand curves are linear, i.e., of the form $p = b - aq$.

Competitive entry at a distance, such as the location halfway between the production center and an economically peripheral point in the market, would then reduce the sales of the firm(s) at the production center by at least 25 percent per firm.[4] In contrast, the location of a new firm at the production center would elicit varying impacts on the earlier firms' sales, depending upon the number of firms at the center. Under Cournot's model, for example, total sales in an n firm market are $n/(n+1)$ of the demand abscissa intercept value b/a, and accordingly the sales of any one of the firms would be $(b/a)/(n + 1)$. Thus, if there was only one seller in the market originally, i.e., $n = 1$, the second seller would de-

4. The 25 percent figure (at the least) stems naturally and simply from the assumptions of zero costs, linear demand, and location at the halfway distant point along a monopolist's line market. The zero cost, linear demand conditions require linear sales totals such that the delivered price at the *natural boundary limit point* of the market equals the b intercept value, hence yielding zero sales. Sales increase linearly the closer the buyer's location is to the seller, in the process forming a demand (sales) triangle. Competitive location at the halfway point along the line market signifies in turn that if the distant firm, in effect, took away from the firm(s) at the production center only the buyers located at his location and those located at points farther away from the production center than his own location, the sales triangle he would have obtained (as carved within the original sales triangle) would be equal in area to $(\frac{1}{2})(\frac{1}{2})$ that of the area of the original sales triangle. Hence his impact on the sales of the firm(s) at the original production center is at least 25 percent.

crease the original seller's sales by 33 percent; in turn, a third seller would have a 25 percent impact on the individual sales of each of two sellers initially in the market, and a fourth seller would have a 20 percent impact, . . . , a tenth seller an 11 percent impact, *ad infinitum*.[5] Clearly, when a very large number of sellers are located together (which possibility depicts a nonspatial market type), the impact of a new entrant at the same site approaches zero. But we have already noted the sales impact of a distant entrant is, at least, 25 percent. Indeed, regardless of the number of firms at the production center, the sales impact is always at least 25 percent along the subject transport route.

The exact percentage impact of a competitive entry at a distance depends on the firm's sales policies with respect to buyers located away from its site in the direction of the production center. But a more important matter than particular percentages is the fact that the very existence of a feasible location at a distance means, *ceteris paribus,* that the firm locating there *could win any price war* with respect to the buyers falling naturally within its market area [4]. Thus the *distant firm attains an oligopolistic position vis-à-vis the firms at the production center.* When distance is costly, pure competition could eventuate if and only if either the buyers or sellers or both are located essentially at the same point in economic space. Under less extreme assumptions, economic space results in an unorganized oligopoly, with the location of a distant firm, and hence the firm itself, always being recognizable. This recognition is part and parcel of the oligopoly markets which characterize the space economy.

Closely related to our basic proposition is the matter of "feasibility of locating at a distance from the production center." In particular, what are the cost relationships that are necessary before a firm would even consider the feasibility of selecting a plant site at some significant economic distance from an established production point? This question is an easy one to answer. In fact the question lends itself to a "general" answer, since the cost relationships which warrant consideration of (and hence possible location at) distant places will be seen to apply regardless of the price system the firm pursues (i.e., regardless of whether it practices discriminatory or nondiscriminatory f.o.b. mill pricing, collusive pricing of a basing point or similar order, etc.). This *general* cost feasibility requirement is simply that any production cost disadvantage at the distant site must be offset by a freight cost advantage on shipments to buyers located at the distant site and at sites still farther removed from

5. Since the sales of any one of the n firms is $(b/a)/(n+1)$, an additional entrant will decrease individual sales by $[(b/a)/(n+1)] - [(b/a)/(n+2)]$. Consequently, the rate of decrease in individual sales as a result of the entry of a new seller is $\{[(b/a)/(n+1)] - [(b/a)/(n+2)]\}/\{[(b/a)/(n+1)]\} = 1/(n+2)$, where $n = 1, 2, \ldots$.

the production center. Marginal production cost plus freight cost from the production center to the distant site must, therefore, exceed (or in certain special cases only equal) the marginal production cost at the distant site. The distant firm under such cost relationship is able to dominate market-area points nearest to its site, particularly in directions still farther from the production center. Such location then is feasible. Most vital, being feasible signifies that the distant firm could win any price war that might arise between it and the firms at the production center, *ceteris paribus*. In turn, oligopolistic relationships arise as the firms at the production center tend to be able to recognize the distant firm's advantage with respect to "very distant buyers." They recognize, at the same moment, the impropriety of price wars designed to maintain sales to these buyers. Spatial competition is coterminous with spatial interdependence, and the oligopolistic market (competitive or collusive) is accordingly a characteristic market form of the space economy.

F.o.b. pricing impacts of new location at a distance: summary and conclusions of distant-entry impacts

A location distant from an existing production point is feasible if and only if any production-cost disadvantage that may hold for the distant site is more than offset by freight-cost savings over the market area surrounding that site. Moreover, demand over this region must be sufficient to warrant establishment of the new firm. Under these preconditions the distant firm may compete actively in price for markets extending from its site towards the older production point, while monopolizing those in the other direction. Any price war could be won by the distant firm, given the required preconditions for its location in the first place. The pricing implication of this conclusion is that oligopolistic relations and a kinked demand curve tend to emerge. The kinked demand curve is therefore brought about by the very fact of distance, and this in turn requires some delivered-cost advantages for each firm distributed over the space. Prices below the kink reflect rival reactions such that any price reduction below the kink elicits price cutting by distant rivals; in an opposite pattern, prices above the kink would be reflective of monopolistic organization.

The abandonment of smooth average revenue schedules does not imply corresponding determination of price at a kink. The kink tends instead to be formed after the price has been determined. Nevertheless, the spatial-kinked demand curve *is not simply a rationalization* which accounts for the stability of prices, as some economists would tend to claim. Rather, through reaction functions, it provides insight into how the initial prices prevalent in the space economy tend to stay fixed for a

while, as well as insight as to why prices resulting from competitive entry at distant locations may become somewhat lower or even higher than they were originally, depending in part upon the shape of the demand curve as affected by (alternative) rival reactions [5].

Impacts of new location at a production point: summary and conclusions of entry at the same point in economic space

The emphasis in this chapter on competition at a distance (i.e., seller dispersion) should not obscure the possibility (and likelihood) of others locating in the same general place as another seller or other sellers (i.e., sellers at a point). The real spatial market is marked typically by a combination pattern of competition at a point and at a distance. Accordingly, as is intrinsic to Cournot's model, one may simply *suggest* that sellers who are located side by side tend – in the absence of collusion, tacit or otherwise – to increase total output and to cause prices to fall. But what about distant entry, or even more particularly the price characteristics of the space economy in general?

Impacts of new locations in general: summary and conclusion

The short-run competition in output which causes prices to fall occurs chiefly in the space economy when new firms locate at given production points. This is not to say that significant entry at a distance (i.e., a large number of new entrants locating at diverse places in the space) will not also cause a breakdown of oligopolistic rigidities. It is, however, our objective only to *stress* the price-lowering impact of oligopolistic *competition* resulting from proximate (identical) locations *and* the short-run oligopolistic price rigidities resulting essentially from the location of spatially separated firms. Manifestly, when many firms enter in diverse directions over economic space, substantial (competitive) output and price impacts must also be expected. (See appendix I to chapter 8.)

IV. *From the short run to the long run*

The short run

One market area often overlaps another in the space economy, and on all sides some compression tends to occur as a result of competitive entry at distant sites. Competition also takes place within a production center's market area. It is essentially the outside-market competitive impact which leads to oligopolistic market patterns, though in fact simpler market relations might never have existed at all anyway.

Sellers locating at a distance may dominate (i.e., control) the buyers located nearest their site. And though the firms at the industrial center might proceed by pricing competitively, competition directed against the distant firm will be fruitless if the distant location is economic. Alternatively, the firms at the center may surrender the distant market, and, as additional distant sites prove profitable, some relocation will tend to occur. With firms located in the outlying districts (in the new market areas) trying to maintain prices at a high level with respect to the production center, the firms concentrated in the production center are provided with a beginning picture of a more profitable economic policy than the classical picture of competition between atomistic firms. As the price at the production center is changed, the distant seller raises or lowers his price, *following price leads* to his advantage and making sure that at least the minimum required size of market area continues to be within his control. In time, the distant seller or, even more likely, the old production center, may gain a position of leadership. It is unlikely, in any case, that atomistic price competition will characterize the old center while strictly complex price patterns exist between it and the new. It is more likely that a complex mix will prevail throughout the market. Though space may not eliminate per se the simple market, the simple perfectly flexible market price system is less likely to exist in an economy in which the function of distance is important.[6] A chainlike relation among firms eventuates on all levels, extending from manufacturing to retailing. It may often lead to *organized oligopoly*, a market pattern which need not concern us further in this book.

The zero windfall competitive oligopoly: the long run

Consider the theory of oligopoly which holds that if free entry prevails along with an absence of collusion among sellers (i.e., no effective organization of oligopolists exists at all), the returns to the factors of production, including the entrepreneur, will be commensurate with classical principles of supply and demand. (And see Appendix II to chap. 10 infra.) The free entry–no collusion requirement must yield payments commensurate with the differentials in skills, risks, and/or uncertainties possessed by or accepted by the factors of production. In effect, a zero windfall applies to the oligopoly. But what are the oligopolist's prices? In fact, what would be his prices in a spaceless market – say, a market in which transport costs on finished products are zero, or one in which all buyers or all sellers are located at a point? In turn, what would be his

6. Conceivably, we could locate people and businesses upward in space rather than horizontally and thereby have a very large number of perfectly competitive firms at a point in economic space. But even here, going up and down would be costly.

prices under the spatial oligopoly situation where sellers and buyers are scattered over the space, and sellers compete with one another for selected buyers? Answers to these questions are sketched below.

V. *The spaceless oligopoly*

Recall from Section II the price equations for the representative firm. Assume there are t rival firms in the market now. Convert the price equation for firm r to total revenue terms and then take the partial derivative with respect to q_r. This yields marginal revenue. Then set marginal revenue equal to marginal costs (assumed to be positive), as in (7-13). In addition, require total revenue to be equal to total costs, as in (7-14). Thus:

$$\frac{\partial R}{\partial q_r} = a_0 - 2b_0q_r - c_0 \sum_{s=1}^{t} q_s = C'(q_r), \tag{7-13}$$

$$\mathrm{TR} \equiv a_0q_r - b_0q_r^2 - c_0q_r \sum_{s=1}^{t} q_s = C(q_r) \equiv \mathrm{TC}, \tag{7-14}$$

where $c_0 = b_0$ in deference to the assumption of homogeneity. Include further in the concept of costs, i.e., both in (7-13) and (7-14), the return required for behavioral uncertainty in the oligopoly market. Furthermore, assume that all firms (and products) are homogeneous so that q_r and t are the unknowns.[7] Solving (7-13) and (7-14) simultaneously would then yield a zero-profit spaceless oligopoly solution of the Cournot type.

VI. *The spatial oligopoly*

It was observed at the end of Section II of this chapter that if f firms are located at a production center, and a given rival firm r is located at a

7. In somewhat greater detail, observe that the assumption all firms are homogeneous would make $q_s = q_r$ in equilibrium; hence, the unknowns in the subsequent equations become q_r and t. However, since t is an integer, these equations may not be solvable in a strict sense. In such a case, (7-14) may be changed to the inequality that follows:

$$a_0q_r - b_0q_r^2 - c_0q_r \sum_{s=1}^{t} q_s \geqslant C(q_r). \tag{7-14$'$}$$

And our problem becomes that of choosing the maximum t which satisfies (7-14) and (7-13).

distance, each new entry at the center (or let us now add at any other competitive location) has the effect of increasing supply (at that time or later) provided the oligopolists are competitive and entry and exit remain free. In effect, the value of c_1 in (7-3)″ falls, and the reaction function of all other firms when aggregated is flatter than it was previously. In other words, p_r, q_r, Π_r will fall. Equations (7-15) and (7-16) can be seen ultimately to apply to the f.o.b. mill pricing firm, where $\Sigma_{s=1}^{\prime}$ is converted to its alternative form $t(c_0)q_s$, while a_0 is rewritten as a_{00} in recognition of the freight-cost impact on the profit-maximizing f.o.b. mill price. More specifically, as is implicit to derivations in Section III of this chapter, a_{00} is smaller than a_0 by a freight-cost subtraction from the demand that is otherwise visualized by firm r.[8] Thus:

$$\frac{\partial R}{\partial q_r} = a_{00} - 2b_0 q_r - t(c_0)q_s = C'(q_r); \tag{7-15}$$

$$\mathrm{TR} \equiv a_{00}q_r - b_0 q_r^2 - t(c_0)q_s q_r = C(q_r) \equiv \mathrm{TC}. \tag{7-16}$$

The f.o.b. mill price of spatial oligopolists is clearly counterpart to the nondiscriminatory price of spaceless oligopolists. The efficiency, want-satisfying, and locational properties of the space economy are covered in [5] and sketched partly in Appendix II to Chapter 8.

VII. *Conclusion*

Competitive impacts on spatial prices have scarcely been touched on in the inquiries of this chapter. At the most, general relationships and views of spatial prices under competition have been set forth. Thus it is that Chapter 8 must examine more precisely the numerical difference between a_0 and a_{00}, doing this in a way corresponding to the inquiries of Chapter 6. In particular, Chapter 8 sets forth the counterpart competitive model to the spatial monopoly model of Chapter 6. Differences between f.o.b. mill pricing *and* discriminatory pricing under conditions of competition in economic space thus serve as part of the subject matter of Chapter 8, as do long-run and short-run impacts of price discrimination.

8. For example, if gross demand is linear, we know from Chapter 6 that the mill price of the spatial monopolist is $b/3$ vis-à-vis $b/2$ for the spaceless monopolist, *ceteris paribus*. Pursuant to the rules of competition, the b value conception changes as the entry of rivals takes place. Equations (7-13) through (7-16) reflect this conception via the parameters a_0 and a_{00}.

References

[1] E. H. Chamberlin. *The Theory of Monopolistic Competition.* Cambridge, Mass., 1946.
[2] A. A. Cournot. *Researches Into the Mathematical Principles of the Theory of Wealth.* Translated by N. T. Bacon. 1897.
[3] M. L. Greenhut. "Effects of Excise Tax and Freight Cost: A Geometrical Clarification." *Southern Economic Journal,* 1955, 21:330–35.
[4] M. L. Greenhut. *Plant Location in Theory and in Practice.* Chapel Hill, N.C., 1956.
[5] M. L. Greenhut, M. Hwang, and H. Ohta. "Observations on the Shape and Relevance of the Spatial Demand Function." *Econometrica,* forthcoming, 1975.
[6] E. M. Hoover. "Spatial Price Discrimination." *Review of Economic Studies,* 1937, 4:182–91.
[7] J. Robinson. *The Economics of Imperfect Competition.* London, 1933.

8. Nondiscriminatory and discriminatory pricing under conditions of free entry in economic space

I. *Introduction*

Chapter 7 constrained our analysis and model of spatial competition to the case of simple nondiscriminatory pricing. But would the competitive firm in economic space, unlike the classical competitive firm, discriminate in price? Would it follow in the footsteps of spatial monopolists and also discriminate against nearer buyers when confronted by competition for distant buyers by sellers located more proximate to these distant buyers? If not, why not? If so, why? These questions must be answered. To answer them, the classical price models of spatial monopoly are recast towards the end of identifying the features of the model which are applicable to the case of spatial competition. The final problem of this chapter will, then, be to determine the long-run impacts of competitive entry on the pricing practice of the spatial firm.

This chapter is based on certain structural and behavioral assumptions: (1) buyers or local markets are evenly distributed along a line, (2) sellers are each distributed discretely along the line, and (3) each seller considers his maximum delivered price to be parametrically given to him as a consequence of the prices and reactions of a rival or rivals. These assumptions lead to the conclusion that the price pattern of the (discriminatory or nondiscriminatory) firm depends ultimately on the degree of spatial competition required for the elimination of all surplus returns in the market, *ceteris paribus*.[1] The spatial competition thus conceived will be found to depend ultimately on the cost-demand-distance relations in the space. Under this framework, a "switching point" or "switching distance" or "limited market size" will be found to exist, a

1. The *ceteris paribus* qualification relates to the model's abstraction from special forces which would determine per se the character and form of the discrimination over space. For example, public opinion could determine the price discrimination pattern of the firm, or a commitment to eliminate a certain rival could elicit a type of spatial price discrimination quite different from that derived from the models used herein. Forces of this kind, to repeat, were omitted from the model in Chapter 6, and they are ignored generally throughout the book.

size short of which the firms price f.o.b. mill, and beyond which they price discriminatorily. By use of diagrams and related illustrations, the several price-distance relations characterizing competition in economic space will be viewed.

It warrants stress that the above assumptions help shed light on some basic aspects of spatial pricing. A more general set with respect to buyer and seller distributions over the space economy could easily be postulated along with a Cournot type of behavioral assumption on rivals' supply (not price). Such a spatially generalized Cournot model will be analyzed in Appendix I to this chapter.

II. *Alternative behavioral assumptions*

Chapter 5 established the proposition that a spatial monopolist necessarily absorbs freight. In Chapter 6 the discriminatory delivered-price schedule was diagramed. The question follows: What is the fundamental impact of a competitive entry at a distance when such entry infringes on the market area of the former monopolist?

One might expect in answer that an upper limit (or ceiling) applies to the subject firm's delivered price schedule as a result of new entry, say at point Q in Fig. 8-1. In fact, it might be expected that the initial seller would find his price schedule *pp* chopped off by a rival's price schedule *p'p'* at some point R on the line market PQ. It is assumed for simplicity that the subject contraction does not affect the firm's marginal

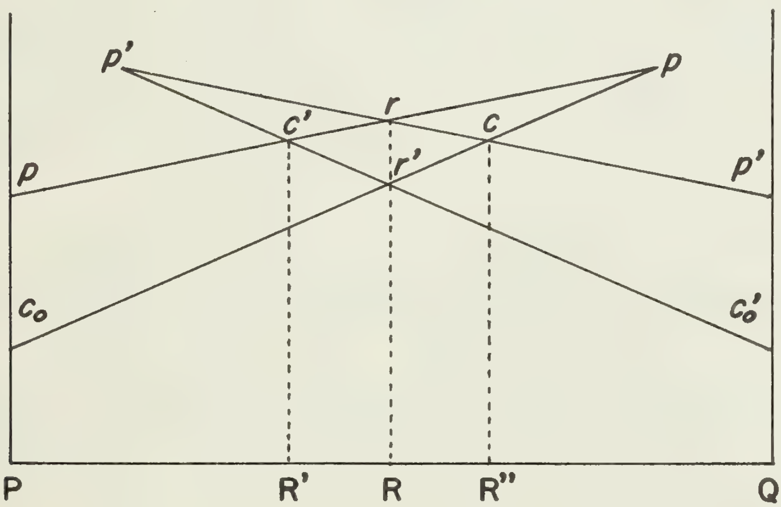


Fig. 8-1 Spatial price schedule under free entry.

cost c_0, the *height* of the freight schedule c_0p, nor *that* of the delivered price *pp*.[2]

The price schedules of the subject and rival firms, given their respective locations, are therefore assumed in Fig. 8-1 to be *pr* and *p′r* respectively. However, if the subject seller *expects* his rival's price schedule to be fixed (and unaltered), for example at *p′p′*, would he not try to capture a part of the rival's market, namely RR″, by quoting as his delivered price a slightly lower schedule than that of his rival over the relevant portion of the space, namely, the price *rc*?[3] If this happens, retaliation by the rival is likely, and the result of the price war will be the appearance of competitive Hooverian price schedules, namely *pc′r′* and *p′cr′* for the two sellers.[4]

Given a particular degree of entry, or more precisely, given the locations of the two firms, the price schedules *pr* and *p′r* will prove to be more profitable than schedules *pc′r′* or *p′cr′*. There is, accordingly, no compelling economic reason why the firms should compete with each other to such an extent that these competitive Hooverian price schedules appear. Only psychological "reactions" would prompt such schedules, derived as they are from excessive competition stemming from the erroneous expectation that a rival's price schedule is unaffected by the price schedule established by a competitive firm.

The relation between two distinct price schedules can be viewed from another side. For this purpose, assume for simplicity that marginal costs are zero and that the buyers' gross demand curves are identical and linear in shape. The unique profit-maximizing mill (net) price for each submarket may easily be determined, as in Chapter 6. The locus of the equilibrium prices for all submarkets are depicted as the line mO in Fig. 8-2(a), where the vertical axis Om_i provides the net price applicable to the ith submarket, while the horizontal axis Oq_i presents the corresponding net demand for the ith market.[5] Line mO is, therefore, a *transform* of *pp* in Fig. 8-1. The transform of price line *pr* in Fig. 8-1 is, in turn, illustrated by the line m_rm in Fig. 8-2(b), where again the two

2. It may be recalled that marginal cost of production c is a function of aggregate output (Chapter 5 above). And since competition reduces the total sales of the firm, the level of c under competition may be different from that under monopoly. Competition may, in other words, shift c_0p and c'_0p' (and correspondingly *pp* and *p′p′*) in Fig. 8-1 upward or downward. For simplicity, however, it is assumed that c itself is constant with respect to aggregate output.

3. Any extension of the market area beyond point R is profitable if the subject firm can ship an increased amount of the good for a price that exceeds the marginal cost of production plus freight cost on the goods. Just so long as the rival's price schedule *p′p′* is unaltered, the subject firm will sell over the portion RR″ in addition to his sales territory PR in order to increase his profits.

4. See Hoover [12, p. 186].

5. Given linear demand curves and zero marginal cost, the equilibrium will be at the halfway point of the net demand curve m_im_i. Then mO is the locus of such equilibrium points. If marginal cost is positive, the locus will lie upward and parallel to mO.

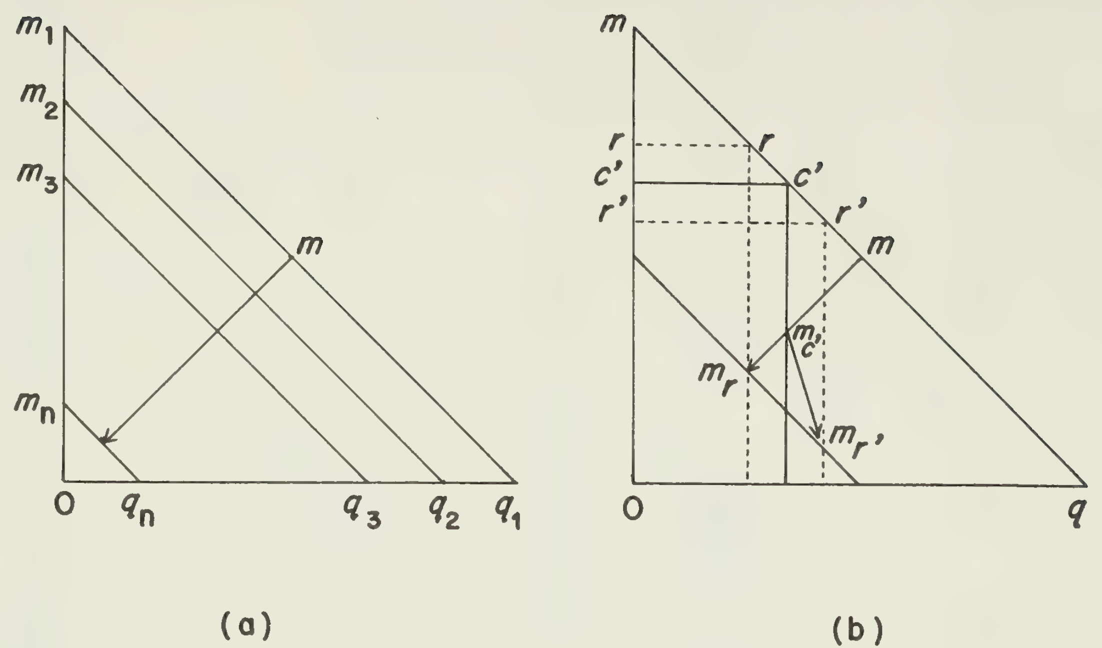


Fig. 8-2 Mill price locus under (a) monopoly and (b) competition.

axes measure the same quantities as those in Fig. 8-2(a). It warrants specification that the maximum (demand effective) price r on the delivered price schedule of the seller is equal to his mill price for the buyer(s) in the furthermost submarket (R in Fig. 8-1) plus the freight rate to that point. The mill price line (locus) in Fig. 8-2(b), therefore, is relevant only over the interval m_rm, as the height of m_r measures the mill price applicable to the furthermost effective market point (R) controlled by the seller; manifestly, the value of the dashed line m_rr stands for the freight rate to that market point. Any market more distant than this one is irrelevant, indeed nonexistent for the subject firm.

The most distant point in the market requires a lower mill price under excessive freight absorption practices (such as are evidenced in the Hooverian price schedule $pc'r'$) than does the most distant point possible under price schedule pr. Thus, the mill price m_r for a buyer or buyers at that point is no longer relevant when the delivered price at R is r' in Fig. 8-1. It falls, instead, from m_r in Fig. 8-2b to $m_{r'}$. More inclusively, the transform of the Hooverian price line $pc'r'$ in Fig. 8-1 is given as $m_{r'}m_{c'}m$ in Fig. 8-2(b). It should be manifest that the locus of the mill price $m_{c'}m$ overlaps with the m_rm schedule, where the value $m_{c'}$ stands for the mill price applicable to the submarket R′; this mill price and the freight rate to such a point sum to the maximum delivered price c' under the Hooverian price discrimination schedule. The mill prices for the submarkets located between points R′ and R, in Fig. 8-1, under the Hooverian price schedule $c'r'$, are clearly less than those corresponding to the price line $c'r$. Moreover, since the mill price for the farthest market at R plus the freight rate to that point must be equal to the *given* delivered price r', the equilibrium for this market extremity is determined at $m_{r'}$ in Fig. 8-2(b), as noted previously. The locus of the mill price bends backward over ranges of excessive price discrimination.[6]

The delivered-price schedule pr in Fig. 8-1 which establishes the mill price locus m_rm of Fig. 8-2(b) provides the firm's maximum profits, given its rival's location. Any deviation from this locus implies a failure to maximize profits. When the locations of the two firms are given, the delivered-price schedules pr and $p'r$ in Fig. 8-1 yield the profit-maximizing mill prices for each firm.

It has thus far been implicitly assumed that the two firms are homogeneous except for location. Relaxing this assumption, for example by proposing unequal marginal costs of production, would not change the above conclusions.[7] We may therefore generally disregard the Hooverian

6. If the delivered price did not decline over the interval R′R, the mill price locus would not bend backward. Instead, it would become vertical if a constant delivered price was established between R and R′, a result traceable to the fact that mill price must decline by the full amount of the relevant freight cost.

7. To prove this, it suffices to show that the market boundary is the same under excessive freight absorption as under, let us call it, normal freight absorption. Thus, suppose (i) that

type of competition and assume that the type of competition which determines the maximum delivered price and the corresponding market boundary is the *pr* schedule of Fig. 8-1. In other words, the firm visualizes a given maximum delivered price when a distant rival firm competes with it for selected buyers. Moreover, under competitive conditions the maximum delivered price will be lower the nearer that the rival is located, i.e., the greater the number of firms in the industry.[8] Parametric treatment of the maximum *delivered* price with a correspondingly obtained market boundary thus helps characterize the world of spatial competition.[9]

III. *Alternative pricing practices under competition*

As more firms enter the industry and as the maximum delivered price accordingly falls, the spatial oligopolist (or duopolist) continues for a while to price like the monopolist (see Chapter 7). His outlook differs

---

the two firms are located d miles apart, (ii) that marginal cost of production is c_0 for firm I and c'_0 for firm II, and (iii) that the freight rate is unitary and the cost of shipping to the market boundary of the initial firm is K. Then the Hooverian boundary is given by the intersection of the following two equations, where each represents the marginal cost + freight cost of the firm:

$$\begin{array}{ll} \text{I} & p = c_0 + K, \\ \text{II} & p = c'_0 + d - K, \\ \therefore & K = \dfrac{c'_0 - c_0 + d}{2}. \end{array}$$

In turn, the normal boundary is determined by the following two equations representing the delivered price schedule of the two firms.

$$\begin{array}{ll} \text{III} & p = \dfrac{b + c_0}{2} + \dfrac{1}{2}K, \\ \text{IV} & p = \dfrac{b + c'_0}{2} + \dfrac{d}{2} - \dfrac{1}{2}K, \\ \therefore & K = \dfrac{c'_0 - c_0 + d}{2}. \end{array}$$

8. Greenhut [11, chap. 13].

9. There exists a third form in which one may introduce spatial competition. Indeed, Mills and Lav [19] and Beckmann [1] actually treated the market boundary (not the maximum delivered price) as parametrically given under f.o.b. mill pricing *and* they implied that their models were of competitive order. However, this kind of treatment is derivable basically from Greenhut's finite distribution of buyers where the spatial monopolist extends his market area [9] and accordingly yields results which do not measure competitive forces perfectly. In fact, if market size is taken as a parameter, mill price increases as the maximum delivered price goes down. Hence a system including parametrically fixed market size is misleading and will not be used herein. Instead, only the maximum delivered price will be taken as a parameter, such that mill price and market size become endogenous variables whose values are to be determined.

from the monopolist strictly because he is subject to an added constraint, namely, his maximum delivered price is parametrically given to him, being less than the price intercept value viewed by the spatial monopolist.

In order to conceive of the competitive spatial oligopoly world in the eyes of a given seller, the two alternative pricing models of Chapter 6 must be converted to the form required by the competitive case. These two alternative models are repeated below. The constraints specified later, i.e., after equation (8-7), will distinguish the competitive equations set from the monopoly equations set:

MODEL I

$$p = f(q); \tag{8-1}$$

$$x = K; \tag{8-2}$$

$$\frac{dp}{dq} q + p = x. \tag{8-3}$$

These equations are subject now to $p \leqslant p'$, where henceforth a prime affixed to a variable such as p will indicate the existence of competition, and where $p' < b$, with b again standing for the price intercept. (As mentioned in Chapter 1, the prime often connotes the first derivative as has been, and will continue to be, our practice; distinction between derivative use and competitive meaning will be clear in the text.) Equations (8-1) through (8-3) are, otherwise, supported by the same simplifying assumptions as those in Chapter 6; hence, they appear exactly the same as equations (6-1) through (6-3).

Model II also consists of the same basic equations as (6-4) through (6-7), again subject to the competitive constraint of p' being less than the price intercept value b. Thus:

MODEL II

$$p = f(q); \tag{8-4}$$

$$x = K; \tag{8-5}$$

$$p = m + x; \tag{8-6}$$

$$\text{Max } \Pi = [m \int_0^{K'_0} f^{-1} (m + K)\, dK]. \tag{8-7}$$

Again, $p \leqslant p'$, where $p' < b$.

The significance of a maximum delivered price for a firm using the

discriminatory price schedule over its space has already been discussed. The significance of a price ceiling constraint on the delivered-price schedule of a nondiscriminatory f.o.b. seller must, however, be discussed further. In this regard, assume the f.o.b. mill seller's objective is also to maximize profits, albeit under nondiscriminatory pricing. The constraint on this seller is that his ("unique") mill price plus freight rate to the market boundary K'_0 must also add up to the exogenously given delivered price p' rather than the price intercept limit $p_0 = b$. It follows that the size K'_0 of his market is a one-to-one function of his mill price m'. Equation (8-7) thus contrasts with (8-3) by containing only *one independent variable,* either the price or distance. Hence, the profit-maximizing condition for the nondiscriminatory competitive firm in economic space can be stated in either variant of (8-7)$'$:

$$\frac{d\Pi}{dm'} = 0 \quad \text{or} \quad \frac{d\Pi}{dK'_0} = 0. \tag{8-7$'$}$$

Determination of the profit-maximizing mill price implies, in other words, determination of the profit-maximizing market size, and vice versa. A parametric treatment of delivered price ceilings involves endogenous treatment of both mill price and market size.

IV. *Selected results*

Total output ($\int q\, dK$) and revenue ($\int pq\, dK$) from the two alternative models of spatial monopoly were derived in Chapter 6. A similar derivation for spatial competition may be set forth, though to simplify the discussion as much as possible, only the results applicable to linear demand will be detailed in the text. As before, each buyer in the market is assumed to have identical gross linear demand curves. However, it will be demonstrated towards the end of the chapter that the conclusions described in this connection hold independently of the shape of the assumed gross demand curve. The present assumption provides numerical results which may be compared with those of Chapter 6. Abstract generalization is thus converted once more to specific numerical form.

Consider first the (profit-maximizing) p^* in equation (8-8) which was derived initially in Chapter 6. (See Table 6-1 above and note that the asterisk is used again to indicate that the subject variable takes on its profit-maximizing value for any given type of market):

$$p = p^* = \frac{1}{2}b + \frac{1}{2}K,\ b \geqslant K \geqslant 0. \tag{8-8}$$

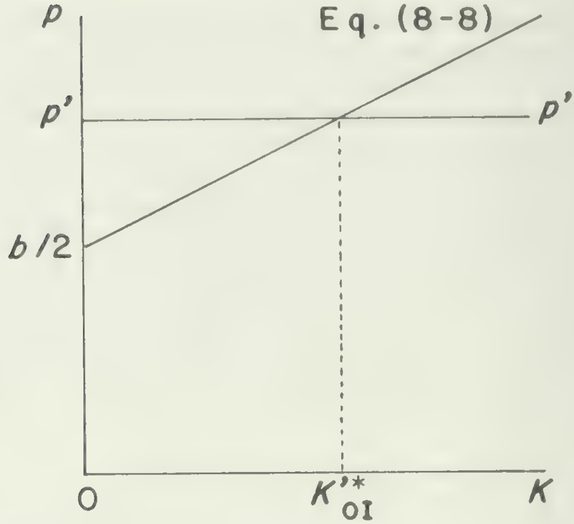


Fig. 8-3 Delivered price ceiling.

Suppose that a delivered price ceiling applies:

$$p = p'b,\ 1 \geqslant p' \geqslant \frac{1}{2}, \tag{8-9}$$

where for computational advantage, we hereafter let $p'b$ (not p' alone) stand for the exogenously given maximum delivered price the seller can charge.[10] Substituting (8-9) into (8-8) provides through (8-10) a constraint that is applicable now to sales distances. This constraint relates in turn to the price restriction provided above. Thus:

$$K'_0{}^* = (2p' - 1)b. \tag{8-10}$$

$K'_0{}^*$ in (8-10) is, in other words, the profit-maximizing size of the market under conditions of spatial competition. Applying a constant price line ($p'p'$) to the price schedule shown in Fig. (8-3) provides a pictorial view of the competitive limit to the firm's trading area. $K'_0{}^*$ under model I —call it $K'_{0\mathrm{I}}{}^*$—has a distance limit value less than b. It is numerically obtained by substituting $K'_0{}^*$ into the integrated output formula and next into the integrated profit functions of Chapter 6. These operations yield (8-11) and (8-12) for model I, again for the case of linear demand:

10. When $p'b$ is at its upper limit b, the system converges to that of spatial monopoly; when it is at its lower limit $b/2$, the system converges to what may be called the spaceless monopoly price because there is no buyer who pays a delivered price greater than the nonspatial price.

$$Q'_{\mathrm{I}}{}^* = \int_0^{K'_0{}^*} qdK = \left(- p'^2 + 2p' - \frac{3}{4}\right) b^2; \tag{8-11}$$

$$\Pi'_{\mathrm{I}}{}^* = \int_0^{K'_0{}^*} (pq - Kq)dK = \frac{1}{12}[1 - 8(1 - p')^3] b^3. \tag{8-12}$$

In model II, the derivation is somewhat more complicated. Pursuant to the assumption of a unit freight rate per unit of distance, we have

$$p = m + K. \tag{8-13}$$

Combining (8-13) with a similar price constraint such as (8-9) yields

$$K'_0 = p'b - m', \; 1 \geqslant p' \geqslant \frac{m'}{b} \geqslant 0. \tag{8-14}$$

Equation (8-14) designates the market size K'_0 under the competitive constraint given parametrically by $p'b$ and an as yet undetermined mill price m'. Taking the derivative of the integrated profit function with respect to distance or mill price, subject to (8-14), and then setting to zero, yields (8-15) or (8-16):[11]

$$m'^* = \frac{1}{3}[2 - \sqrt{(3p'^2 - 6p' + 4)}]b \tag{8-15}$$

$$K'_0{}^* = \frac{1}{3}[3p' - 2 + \sqrt{(3p'^2 - 6p' + 4)}]b \tag{8-16}$$

Fig. 8-4(a) presents the solutions m'^* and $K'_0{}^*$ graphically. The vertical axis measures the delivered price; the horizontal axis depicts the distance from the seller site. Given p' (e.g., as $p'p'$ in the figure), the value of $K'_0{}^*$ for model II is determinable (i.e., as $K'_{0\mathrm{II}}{}^*$) via (8-16). The curve $K'_{0\mathrm{II}}{}^*K'_{0\mathrm{II}}{}^*$ in Fig. 8-4(a) is thus the locus of the maximum distance in terms of a parametrically given maximum delivered price $p'b$.[12] Given p' and $K'_0{}^*$, the mill price is determinable via (8-14). And the curve $m'_{\mathrm{II}}{}^*m'_{\mathrm{II}}{}^*$ is, therefore, the locus of the equilibrium mill price in terms of the maximum sales distance $K'_0{}^*$, where $K'_0{}^*$ in turn is determined by $p'b$ in (8-16). Fig. 8-4(a) demonstrates the principle that spatial competition under model II *reduces the mill price* as well as the firm's market size. Thus, the delivered-price schedule over space under model II *shifts*

11. Either of the two is derivable from the other, taking into consideration equation (8-14).

12. The line $K'_{0\mathrm{II}}{}^*K'_{0\mathrm{II}}{}^*$ is monotonically increasing for the domain given in (8-14), i.e., $dK'_0{}^*/dp' > 0, \; \forall p' \ni 1 \geqslant p' \geqslant 0$.

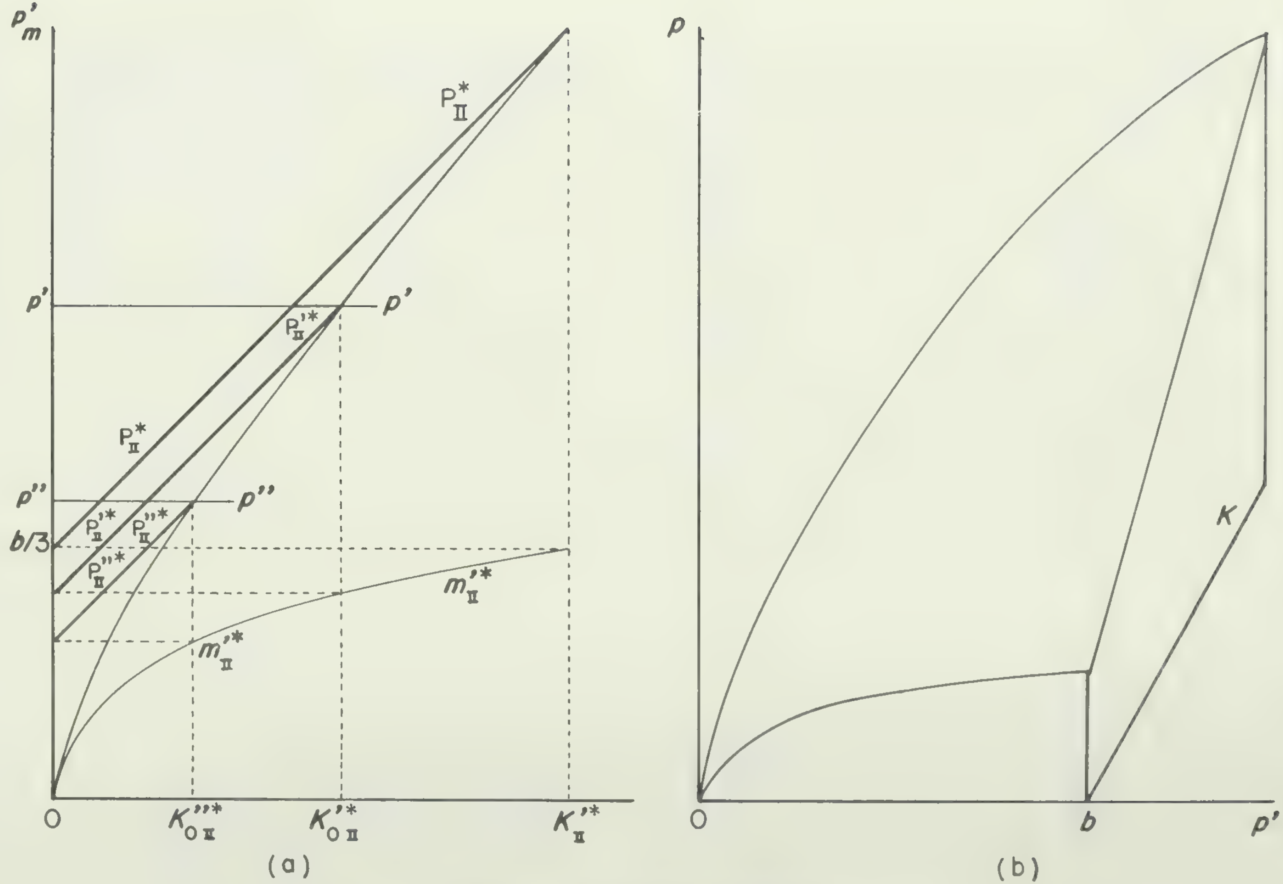


Fig. 8-4 Impacts of competition on prices and market size.

downward and shrinks in size (e.g., compare $P_{\mathrm{II}}{}^*P_{\mathrm{II}}{}^*$ with $P'_{\mathrm{II}}{}^*P'_{\mathrm{II}}{}^*$ and $P''_{\mathrm{II}}{}^*P''_{\mathrm{II}}{}^*$).

The relations illustrated in Fig. 8-4(a) are basically three-dimensional and therefore can be transformed to a cubic relation with three axes as in Fig. 8-4(b). In the figure, the vertical p axis measures the delivered price while the horizontal p' axis and K axis stand for the parametric maximum delivered price and distance, respectively. This solid diagram is also designed to show that a lowering of p' as a result of competition will lower mill price, i.e., the value of p at $K=0$. This, in turn, will reduce the size of market, i.e., the limiting (maximum) value of K. In effect, the diagram, therefore, indicates the delivered price over any distance K under the given degree of competition p'.

It warrants mention that the maximum sales distance is greater under model I than under model II when spatial competition is on a relatively small scale, but the opposite holds when the competition is comparatively substantial. This relation is illustrated in Fig. 8-5(a), which is a simple combination of the relevant parts of Figs. 8-4(a) and 8-3. A similar relationship holds also for the mill price. Thus, a weaker degree of spatial competition, or no spatial competition at all, implies a generally lower mill price under model I than under model II. A greater amount of competition carries the opposite effect. We see, accordingly, in Fig. 8-5(b) that $m'_{\mathrm{I}}{}^*$ and $m'_{\mathrm{II}}{}^*$ represent the mill prices applicable to the respective market boundary points $K'_{0\mathrm{I}}{}^*$ and $K'_{0\mathrm{II}}{}^*$, given in turn by the relevant degree of competition p'. It must be noted that the unique value of $m'_{\mathrm{II}}{}^*$ that relates to p' is also applicable to any point nearer to the seller than the market boundary $K'_{0\mathrm{II}}{}^*$; correspondingly, the value of $m'_{\mathrm{I}}{}^*$ applies to a single point $K'_{0\mathrm{I}}{}^*$, again given p'.

Mill price under model I is a one-to-one function of delivered price p (whose upper limit value is p'), while delivered price p is a one-to-one function of distance K (whose upper limit value is $K'_{0\mathrm{I}}{}^*$). Mill price under model II is, however, only a one-to-one function of the parametrically given maximum delivered price p', which value sets the upper limit to the delivered price p. Finally, observe that the mill price is uniquely given in model II for any p such that $p' \geqslant p \geqslant m'$, while different mill prices apply for different delivered prices under model I.

V. *The switching point analyzed*

Figs. 8-5 reveal "switching" points in mill prices as well as in sales distances. This "relation-set" leads to the question whether there also exist switching points in profits and outputs. An affirmative answer applies which, in turn, has a remarkable implication! *There exists* a

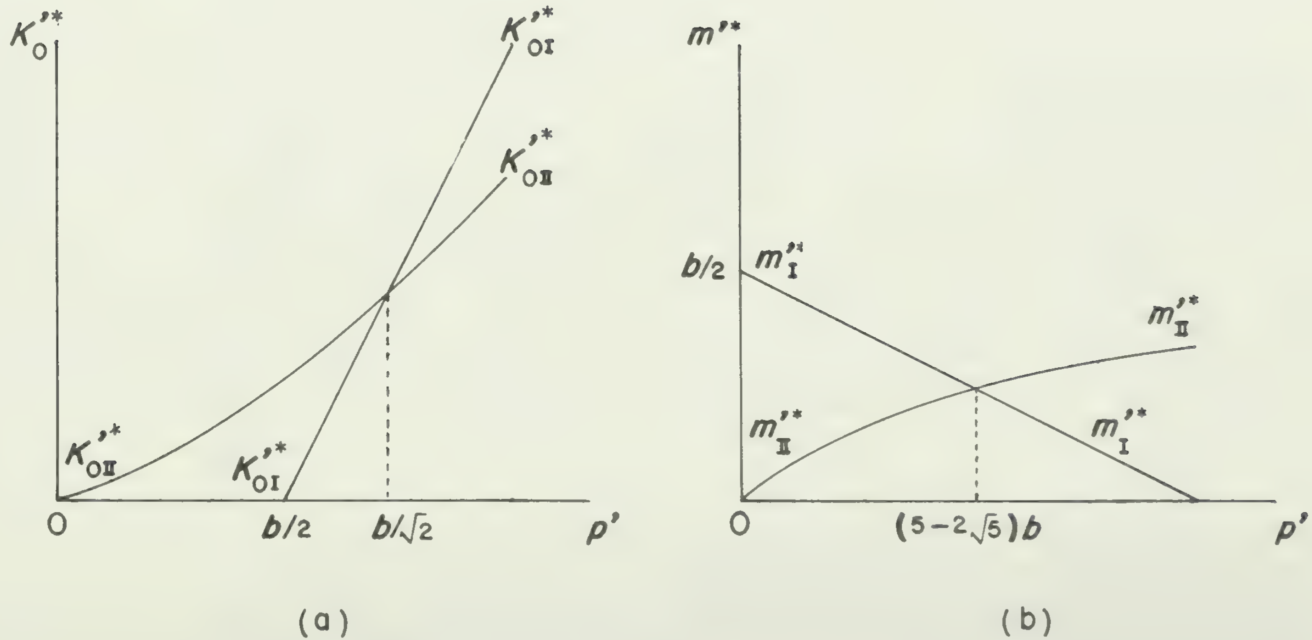


Fig. 8-5 Competitive market size and mill price under alternative pricing practices.

switching point in profits such that a particular pricing technique, viz., the discriminatory one, turns out to be less profitable than the alternative (f.o.b.) pricing system *as the degree of competition increases.* The firm, in other words, switches its profit-maximizing pricing technique when subject to a particular intensity of competition. This condition can be proved as follows.

Substitute (8-16) into the integrated output function of model II. This yields

$$Q'_{\mathrm{II}}{}^{*} = \int_0^{K'_0{}^*} q\, dK = (b - m'^*)K'_0{}^* - \frac{1}{2}K'_0{}^{*2} \tag{8-17}$$

$$= \frac{1}{9}\left[-3p'^2 + 6p' + \sqrt{(3p'^2 - 6p' + 4)} - 2\right]b^2$$

$$\frac{dQ'_{\mathrm{II}}{}^{*}}{dp'} = 0 \quad \text{when } p' = 1, \tag{8-18}$$

$$= \frac{1}{3}(1 - p')\left[2 - \frac{1}{\sqrt{(3p'^2 - 6p' + 4)}}\right]$$

$$> 0, \ \forall\, p' \ni 1 > p' \geqslant 0.$$

It is indicated by (8-18) that as p' falls, the total output of a firm decreases *monotonically* in accordance with the lowering of its maximum delivered prices.[13] Quantity $Q'_{\mathrm{II}}{}^*$ of (8-17) can be compared diagrammatically with $Q'_{\mathrm{I}}{}^*$ of (8-11), as in Fig. 8-6(a).[14] Similarly, since mill price is also a monotonic function of p' under model II [see (8-15)], the firm's profit under model II, i.e., $\Pi'_{\mathrm{II}}{}^* = m'^*Q'_{\mathrm{II}}{}^*$, is also a monotonic function of p' under the relevant domain, i.e., $1 \geqslant p' \geqslant 0$. $\Pi'_{\mathrm{II}}{}^*$ can then be compared with $\Pi'_{\mathrm{I}}{}^*$ in (8-12), as in Fig. 8-6(b) and note 14 above.

There exists, accordingly, a switching point shown by the intersection of the two curves in Figs. 8-5(a) or (b). To the right of the intersection point, the discriminatory pricing technique yields more outputs (profits) than the nondiscriminatory pricing technique. To its left, the opposite

13. $dQ'_{\mathrm{II}}{}^*/dp' \geqslant 0$ applies because $2 \geqslant \sqrt{(3p'^2 - 6p' + 4)} \geqslant 1, \ \forall p' \ni 1 \geqslant p' \geqslant 0$.
14. Differentiate (8-11) with respect to p'. This yields

$$\frac{dQ'_{\mathrm{I}}{}^*}{dp'} = 2(1 - p')b^2 > 0, \quad \forall p' < 1. \tag{8-11)$'$}$$

$Q'_{\mathrm{I}}{}^*$ is, therefore, also a monotonically increasing function of p' over the domain $0 \leqslant p' < 1$. Observe further that pursuant to (8-11), $Q'_{\mathrm{I}}{}^* = Q_{\mathrm{I}}{}^* = (\frac{1}{4})b^2$ when $p' = 1$, while $Q'_{\mathrm{I}}{}^* = 0$ when $p' = \frac{1}{2}$. In turn, (8-17) demonstrates that $Q'_{\mathrm{II}}{}^* = Q_{\mathrm{II}}{}^* = (\frac{2}{9})b^2$ when $p' = 1$, and $Q'_{\mathrm{II}}{}^* = 0$ when $p' = 0$. In sum, $Q'_{\mathrm{I}}{}^* > Q'_{\mathrm{II}}{}^*$ when $p' = 1$, while $Q'_{\mathrm{I}}{}^* = 0$ when $p' = \frac{1}{2}$ and $Q'_{\mathrm{II}}{}^* = 0$ when $p' = 0$. Each Q'^* is a monotone function under its respective domain. These conditions are sufficient for Fig. 8-6(a). A similar statement holds for Π'^* in Fig. 8-6(b), involving in its turn the differentiation of equation (8-12).

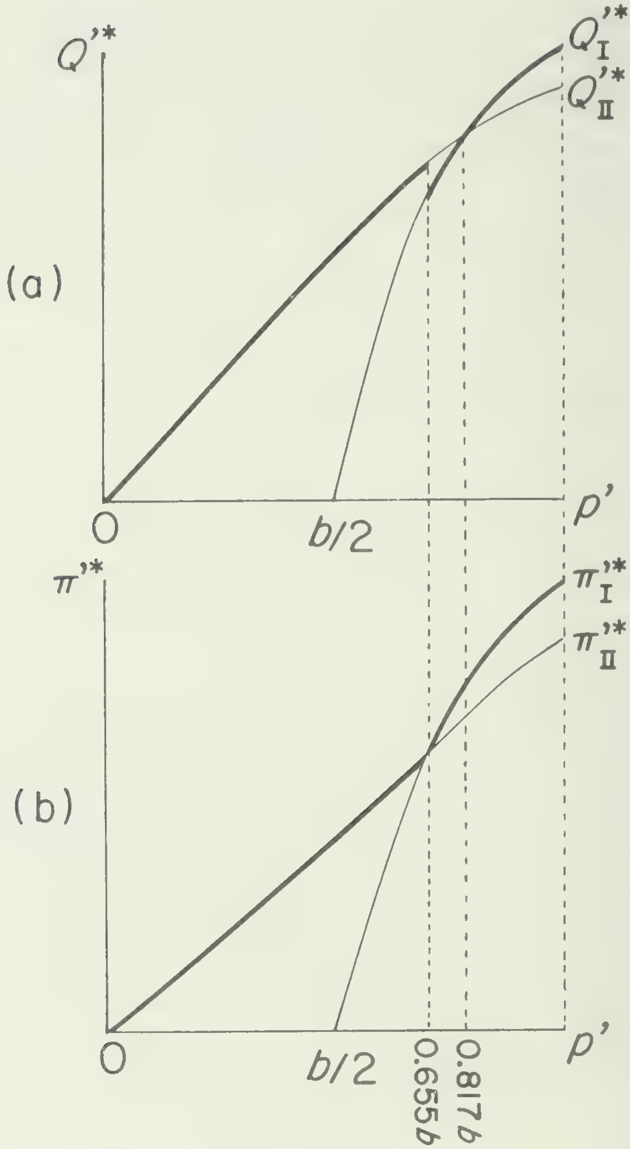


Fig. 8-6 Competitive output and profit under alternative pricing practices.[15]

15. According to the result of our computer runs, the switching point for Π'^* lies to the left of the switching point for Q'^*. This implies that for short distances after the discriminatory policy has become more profitable than the f.o.b. mill policy, output produced under the former policy (I) will continue to be smaller than the output produced under the latter policy (II), i.e., $Q'_I{}^* < Q'_{II}{}^*$ because of the generally higher discriminatory price level. In opposite sequence, as competition lowers p' to the level where f.o.b. pricing is more profitable, $Q'_{II}{}^*$

statement applies. As the degree of competition increases, a spatial seller increases his profits by switching from discriminatory to nondiscriminatory pricing. Moreover, the switching point (i.e., when to switch) can be determined unambiguously if the demand conditions are known (e.g., a linear demand function, as in our present discussion). It must be emphasized that the existence of the switching point is itself independent of the shape of the demand curve for three reasons. (1) Profit obtained under spatial discriminatory monopoly is strictly greater than that obtained under simple monopoly, regardless of the shape of the gross demand curve (Chapter 5). (2) Profits must decline monotonically as p' decreases with increasing competition, again regardless of the shape of the demand curve. And (3) the minimum value of p' related to $\Pi'_{\mathrm{I}}{}^*$ is the mill price applicable to the buyer located at the seller's site. This price, in general, is greater than the marginal cost of production, while the p' related to $\Pi'_{\mathrm{II}}{}^*$ may equal marginal cost. Note that the zero MC assumption inherent in Fig. 8-6 or 7 is not requisite to the present statement. Thus, $p'_{\mathrm{I}} > p'_{\mathrm{II}}$ as Π'^* approaches zero. So any switching point in the profit curves must exist in terms of p' regardless of the form of the demand or cost functions.

VI. *Conclusion*

Do the above results lead to the conclusion that a firm under spatial competition resorts—in general—to the nondiscriminatory f.o.b. pricing technique? To answer this question, let pure (windfall) profit be defined as an excess of profit Π over fixed cost F. Then assume three alternative levels of fixed cost, as illustrated by F_1F_1, F_2F_2, and F_3F_3 in Fig. 8-7. The relevant $\Pi'_{\mathrm{I}}{}^*\Pi'_{\mathrm{II}}{}^*$ curves in Fig. 8-7 are copied in turn from Fig. 8-6(b). If the fixed cost is as high as F_1F_1, model I yields zero profit equilibrium at P_1. At this point, model II yields negative profit. The profit-seeking firm must adopt discriminatory pricing if the fixed cost is higher than F_2F_2, even under the zero-profit competitive equilibrium. If fixed cost is as high as F_2F_2, the zero-profit equilibrium is obtained at P_2. Models I and II are equally profitably at this level of fixed cost, and the firm is indifferent to the subject pricing technique. If, finally, fixed cost is lower than F_2F_2, e.g., as low as F_3F_3, model II yields zero-profit equilibrium at P_3, while model I yields negative profit. The firm must switch its pricing practice from discriminatory to f.o.b. pricing as the degree of competition

---

becomes increasingly greater than $Q'_{\mathrm{I}}{}^*$. As p' is lowered below $p' = 0.6558b$ it follows that $\Pi'_{\mathrm{II}}{}^* > \Pi'_{\mathrm{I}}{}^*$ and $Q'_{\mathrm{II}}{}^* > Q'_{\mathrm{I}}{}^*$. The locus of the Π'^* and Q'^* which the firm chooses under conditions of spatial competition are respectively depicted in the heavy solid lines in this Fig. 8-6.

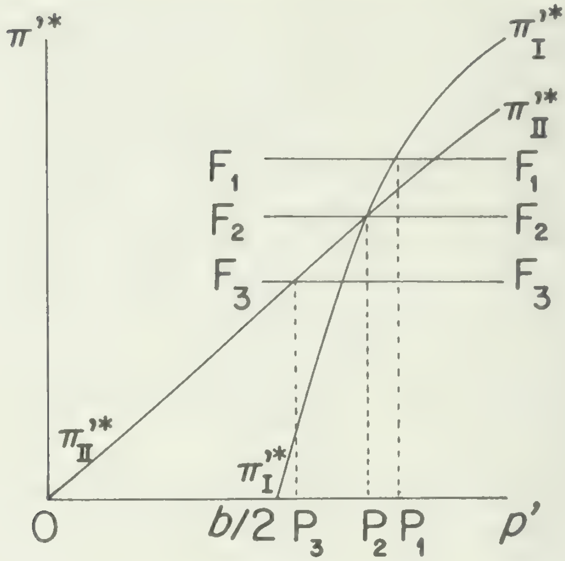


Fig. 8-7 A determinant of pricing practices under competition.

increases, i.e., as the maximum delivered price falls to lower and lower values.

The level of fixed cost plays the decisive role in determining whether or not a firm applies discriminatory or nondiscriminatory pricing under the zero-profit competitive equilibrium. This conclusion is supported by the simplifying assumptions that both the freight rate and the marginal cost of production are fixed. However, it should be stressed that these costs play a role similar to fixed cost. The higher they are, the lower will be the $\Pi'_{I}{}^{*}\,\Pi'_{I}{}^{*}$ and $\Pi'_{II}{}^{*}\,\Pi'_{II}{}^{*}$ curves. In turn, the switching point also moves downward. But then the fixed cost level F_2F_2 requires the firm to discriminate in the zero-profit competitive equilibrium. It follows that F_2F_2 no longer defines the indifference level of costs between the alternative spatial pricing practices under the given conditions of competition. It must be concluded that the relative importance of *costs in general,* i.e., freight as well as production costs, *determines the firm's pricing practices under spatial competition.* It further follows that if, as is claimed elsewhere [11], the friction of distance and/or uncertainty is a significant cost in a space economy, a weak form of competition tends to wipe out economic surpluses before the switching point is reached which favors nondiscriminatory pricing. Thus, a firm may be expected to adopt discriminatory pricing over space even in the presence of spatial competition and the long-run prospect of zero profits. Empirical testing of this proposition should be of interest to many economists in several ways, particu-

larly those concerned with antitrust laws such as the Robinson-Patman Act.

Appendix I: *An alternative model of spatial pricing under competition**

This appendix generalizes Cournot's model so that it can focus attention on certain facets of spatial price discrimination. For this generalization we assume (1) n local submarkets or buying points scattered over a land surface, (2) m sellers also dispersed over the land surface, (3) homogeneous products, and (4) Cournot type of competition, i.e., each firm considers its rival firms' supply to be fixed to any given buying point.

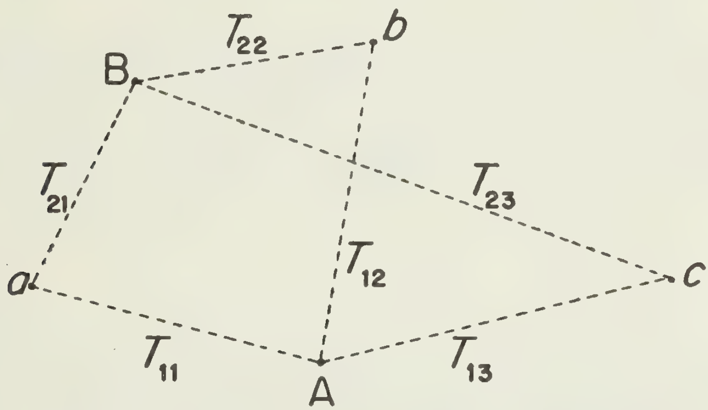


Fig. 8-8 Market points and firm sites distributed over economic space.

The distribution of markets and sellers conceived of herein is illustrated in Fig. 8-8 for the case of two sellers and three buying points, i.e., $m = 2$ and $n = 3$. Distance costs (in terms of freight rates) from firm A to market points a, b, and c are represented by T_{11}, T_{12}, T_{13}, and costs of distance from firm B to the same market points are represented by T_{21}, T_{22}, T_{23}. Average cost of transportation to market point a is defined as the average of all firms' freight costs to that buying point, i.e., $(T_{11} + T_{21})/2$. Correspondingly, average freight costs to market points b and c are respectively $(T_{12} + T_{22})/2$ and $(T_{13} + T_{23})/2$.

* This appendix is based on a recent theoretical development by John Greenhut for publication elsewhere. The authors are indebted to him for his permission to freely reorganize part of his material for use herein.

Define the local demand function at market point j as

$$p_j = f_j(q_{\cdot j}), \quad j = 1, 2, \ldots, n, \tag{8-19}$$

where $q_{\cdot j}$ represents the quantity demanded at price p_j in market j. This quantity is to be distinguished from q_i, which stands for the quantity produced by the ith seller. These variables consist of the partial aggregates of the quantity supplied by each of the i firms selling to market point j, as represented by q_{ij}. Thus we have

$$q_{\cdot j} = \sum_{i=1}^{m} q_{ij}, \quad j = 1, 2, \ldots, n; \tag{8-20}$$

$$q_i = \sum_{j=1}^{n} q_{ij}, \quad i = 1, 2, \ldots, m. \tag{8-21}$$

Equation (8-20) establishes the equilibrium condition for the quantity demanded at point j and the quantity supplied by all m firms; (8-21) states that the quantity produced by the ith firm must equal the total quantity shipped by this firm to its n market points.

Profits for the ith firm are definable as the firm's total revenues on sales to n market points less total production and transportation costs:

$$\Pi_i = \sum_{j=1}^{n} p_j q_{ij} - C_i - \sum_{j=1}^{n} T_{ij} q_{ij}, \quad i = 1, 2, \ldots, m, \tag{8-22}$$

where C_i stands for the cost of production as a function of q_i (which is independent of total q_{ij}) while T_{ij} stands for the freight rate (per unit quantity) applicable to the distance from the firm's site i to buying site j. The problem for firm i is to maximize Π_i subject to Cournot type of competition. Thus (8-23) below applies:[1]

$$\begin{aligned}\frac{\partial \Pi_i}{\partial q_{ij}} &= p_j + f_j'(q_{\cdot j})q_{ij} - c_i - T_{ij}, \quad c_i = C_i' = dC_i/dq_i \\ &= p_j\left(1 - \frac{\alpha_{ij}}{e_j}\right) - c_i - T_{ij} = 0, \quad \forall i, j,\end{aligned} \tag{8-23}$$

where e_j stands for the elasticity of demand at buying point j, α_{ij} is the ith firm's market share at buying point j,[16] and c_i is firm i's marginal cost

16. It is numerically possible that some, if not all, α_{ij} could be negative. This result follows because the marginal costs applicable to some market point for a given firm could exceed the equilibrium price at that point, i.e., $c_i + T_{ij} > p_j$, which implies via (8-23) that $\alpha_{ij} < 0$. Any equation in (8-23) which involves such an economically meaningless condition

of production.[17] It follows from (8-23) that the equilibrium price prevailing at market point j reflects the elasticity of demand at point j, the individual firm's market share at this market point, its marginal cost of production, and its transportation cost to this market point. It goes without saying that the price determinants and other variables applicable to market point j must be determined simultaneously.

In Walrasian style, the number of equations and unknown variables may be counted over equations (8-19) to (8-23). We find n equations in (8-19), n in (8-20), m in (8-21), m in (8-22) and mn in (8-23), i.e., $2(m+n)+mn$ in total; in turn, there are n variables for p_j, n for $q_{\cdot j}$, m for q_i, m for Π_i, and mn for q_{ij}, i.e., the same $2(m+n)+mn$ in total. Thus we may conclude that our model of spatial competition can be solved for all unknown variables.

Operationally meaningful results require, however, summating (8-23) over all i to obtain

$$mp_j + f_j'(q_{\cdot j})q_{\cdot j} - \sum_{i=1}^{m} c_i - \sum_{i=1}^{m} T_{ij} = 0, \quad \forall j. \tag{8-24}$$

Dividing both sides of (8-24) by m yields

$$p_j\left(1 - \frac{1}{me_j}\right) - c_i\dagger - T_{ij}\dagger = 0, \quad \forall j, \tag{8-24$'$}$$

where $c_i\dagger$ and $T_{ij}\dagger$ stand respectively for the average marginal costs of production and transportation, i.e., $c_i\dagger = (1/m)\sum_i c_i$ and $T_{ij}\dagger = (1/m)\sum_i T_{ij}$. Most significantly, the unknown market share variable, α_{ij}, cancels out in the process of aggregating costs and sales over the markets.[18] The price at market point j is, therefore, fully determined by

---

is, therefore, inapplicable and must be disregarded. For simplicity, however, we henceforth assume, unless otherwise stated, that all m firms sell to all n market points, i.e., $c_i + T_{ij} \leqslant p_j$ for all i and j.

17. The subject equation requires the assumption $p_j \geqslant c_i + T_{ij}$ for all i and j. Without this condition, the summation takes place only up to $m_j < m$, where m_j represents the number of firms selling to market point j. This m_j is determined by the number of firms whose particular marginal revenue in that market equals the firm's marginal costs, $c_i + T_{ij}$.

18. The summation of (8-23) can be supported more formally than by just the assumption given in note 17 above. In particular, realize that a stable market price implicitly requires firms to be pricing at the average cost level rather than for low-cost firms to be breaking the equilibrium by undercutting prices at any given market point. Institutional constraints, such as antitrust laws, customer goodwill, etc., also serve to prevent the low-cost firm from setting the price. In fact, the implicit requirement that higher than average cost firms might remain in the market is not so troubling in spatial economics as in nonspatial economics, because maintenance of a foothold in one submarket among many is a likely practice among sellers in a dynamic world. Final justification for taking m parametrically

$$p_j = \frac{c_i\dagger + T_{ij}\dagger}{1 - (1/me_j)}, \quad \forall j, \tag{8-24)''$$

provided $c_i\dagger$ is given as a datum.[19] Although $c_i\dagger$ in general is not a given constant (since $c_i\dagger$ is a function of q_i), analytical simplicity, operational practicality, and reasons provided previously in notes 18 and 19 justify our taking it as a datum for equations (8-24) to (8-24)''.

Equation (8-24)'' uncovers important relationships between the equilibrium price p_j and the determinants of that price. Thus, the greater the number of firms selling to any market point j, and/or the higher the elasticity of demand at that point, the lower is the equilibrium price. This relationship is but a spatially generalized version of Cournot's theorem on nonspatial prices and number of sellers. More important is the relation that the higher the *average* marginal costs are of production and/or *transportation* applicable to market point j, the higher will be the equilibrium price at this point. It follows that the more distant the firms locations are from market point j, the higher are the prices at such market point.

The p_j being determined, the relationship between this variable and the firm's transportation cost T_{ij} is readily specifiable. Call this relationship the individual seller's delivered price schedule, and denote it henceforth as DPS. For any given firm, its DPS may be evaluated by plotting the transportation costs T_{ij} from the plant site of the subject firm i to market point j against the relevant delivered prices p_j, as determined by (8-24)''. It is manifest that the DPS may not be in simple relationship to the firm's transportation costs T_{ij} because the delivered price p_j is a function of $T_{ij}\dagger$, not directly of T_{ij}. For example, even if the firm's transportation cost increases, say from T_{ij} to $T_{i(j+1)}$, the delivered price p_j may remain constant or even decline, depending upon the change (possibly zero or negative) taking place in average transportation cost $T_{ij}\dagger$. F.o.b. pricing is, therefore, in principle inapplicable to a spatially *competitive* equilibrium in our generalized Cournot model. Competition over economic space is a sufficient condition for price discrimination, and the delivered price to any one location may be said to be unaffected by the delivered price to any other location. Evaluation of the slope of DPS, in any case, becomes central to the analysis of competitive impacts on discriminatory pricing in the space economy.

over the entire space involves a Friedman type argument; as will be realized later, if the number of competitors is treated as a variable, the curvature of the price schedule is altered, albeit underlying forces and basic results are unchanged; the extra realism gained *in this instance* (as with Occam's razor) would not be worth the price.

19. This requirement is readily fulfilled because (1) the average transportation cost $T_{ij}\dagger$ is known under parametric treatments of T_{ij}, (2) the number of competing firms m is known, and (3) the elasticity of demand prevailing at market point j is either constant or else a function of the price p_j.

II

Our analysis may be simplified and yet well illustrated by postulating a general polynomial demand functions for all market points, as was done in the Appendix to Chapter 6. Thus, we specify the basic price function:

$$\frac{p_j}{\beta} = 1 - \left(\frac{q_{\cdot j}}{\alpha}\right)^x, \forall j. \tag{8-25}$$

The elasticity of demand is then given by

$$e_j = \frac{1}{x}\left(\frac{p_j}{\beta - p_j}\right), \quad \forall j. \tag{8-26}$$

Substituting this e_j into (8-24)″ in turn yields

$$p_j = \frac{1}{m+x}(x\beta + mc_i\dagger) + \frac{m}{m+x} T_{ij}\dagger, \quad \forall j. \tag{8-27}$$

The derivative of p_j with respect to T_{ij} is

$$\frac{dp_j}{dT_{ij}} = \frac{m}{m+x}\frac{dT_{ij}\dagger}{dT_{ij}}, \quad \forall i. \tag{8-28}$$

On the basis of the simplified model set forth above, the impacts on delivered-price schedules of different sellers' (and buyers') locational arrangements over space can now be determined. If, for example, spaceless demands are identical and all sellers are located at the same (given) production center, each seller is subject to identical transport cost $T_{ij} = T_j$. Delivered price under the defined spatial alignment may be readily viewed in (8-27), where the first term on the right side of (8-27) is the mill price and the coefficient of $T_{ij}\dagger$ gives the slope of the schedule. Because $T_{ij}\dagger$ increases everywhere at the same rate as T_j, the DPS is linear. When $x > 0$, the slope is positive but less than unity; discrimination therefore proceeds in favor of distant buyers.

The effect on delivered prices of the number of competitors at the production center can be obtained by deriving the partial derivatives with respect to the number of competitors of (a) the DPS, (b) the mill price, and (c) the slope S of the DPS; only $\partial S/\partial m > 0$, the other partials being negative. Hence, the greater m is, *ceteris paribus,* the lower are delivered and mill prices, but the steeper is the slope of the DPS. Mill price in fact approaches $T_{ij}\dagger$ *and* the slope of the delivered-price schedule unity as

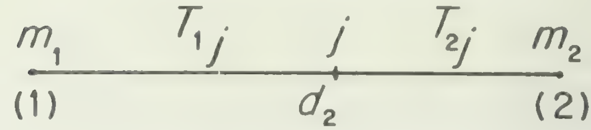


Fig. 8-9 Two production sites on the buyer's line.

m becomes increasingly great. Perfect competition culminates in a price schedule which resembles nondiscriminatory f.o.b. mill pricing, and which in fact becomes the f.o.b. price system.

Another analytical handle to spatial pricing under competitive conditions may be gained by conceiving of sellers located at *two* alternative production centers, each of which can serve a given market space. It is readily apparent that if every seller can supply every buyer located between points (1) and (2) in Fig. 8-9, average transport cost to a buyer who is T_{1j} dollars of distance from site (1) and $T_{2j} = d_2 - T_{1j}$ dollars from site (2) is given by

$$T_{ij}\dagger = \frac{1}{m_1 + m_2} \left[m_1 T_{ij} + m_2 (d_2 - T_{ij})\right], \quad i = 1, 2, \quad j = 1, 2, \ldots, n. \tag{8-29}$$

Utilizing $m = m_1 + m_2$ along with equations (8-27) and (8-29) enables us to derive the delivered price of each seller, the mill price of the sellers located at site (1), the delivered price at site (2), and the slope of the DPS of the sellers located at site (1). The alternative delivered price

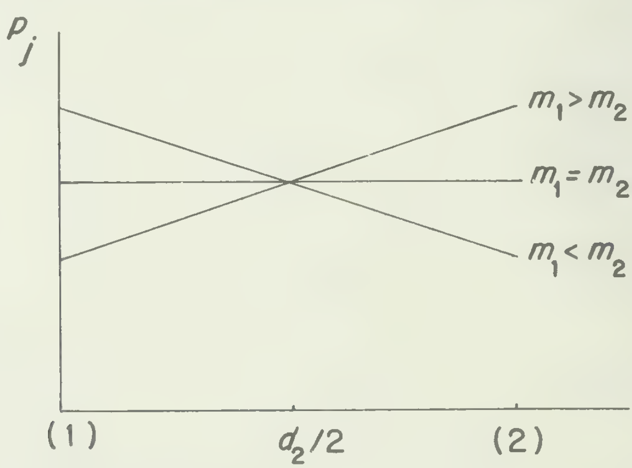


Fig. 8-10 The DPS when buyers are located on a line between two sellers.

schedules given in Fig. 8-10 depict three distinguishable and important schedules based on alternative values of m_1 and m_2.

Another linkage between delivered prices and the effect of a distant production site may be easily obtained by taking the partial of S with respect to m_2, which is negative. The incidence of an increase in m_2 is a decrease in the slope of DPS. A most compelling implication is that *the more competitive the distant location is (i.e., the larger m_2 is), the greater will be the freight absorption rate (given by $1 - S$) of the firms selling at site* (1).[20]

One special result of distant competition can, therefore, be uniform pricing. This rather prevalent price practice in the United States would arise a priori when competition is of the same weight at alternative production centers. Although products advertised nationally are often subject to low freight cost and thereby are readily subject to uniform pricing, our theory indicates that an even distribution of sellers is another compelling force behind the use of uniform price schedules. It may, in fact, be the case that when a product's price is advertised nationally rather than by region, such policy itself is attributable to the fact of even-weighted production centers. A serious limitation to this thesis does, however, exist, as may be seen when a slightly different specification of spatial interdependencies is made. Thus, for example, if the distance separating the centers is so great that the firms in each center do not compete with distant rivals in the neighborhoods of their production center, the DPS over the region proximate to site (1) can easily be shown as

$$p_j = \frac{1}{m_1 + x}(x\beta + m_1 c_i \dagger) + \frac{m_1}{m_1 + x} T_{1j}.^{21} \tag{8-30}$$

Over the competing region, the DPS is given by (8–30) except that the $T_{ij}\dagger$ of (8-29) must be substituted for T_{1j} in (8-30). The delivered-price schedule illustrated in Fig. 8-11 is, therefore, a simple combination of the two schedules. The caveat applies and warrants specification that in constructing Fig. 8-11, m_1 was assumed to be greater than m_2.

The delivered-price schedule of the representative firm located at site (2) may be constructed like that for the firm at site (1). Manifestly, a downward curve in prices begins at the location at which the distant competition from m_1 firms occurs. The gist of the relationship is simply that the DPS is a combination of schedules under varying degrees of

20. $(1 - S)\,T_{ij}$ is the freight absorption, since delivered prices rise by $ST_{ij} = \dfrac{m}{m+x}\,T_{ij}$ while freight cost rises by T_{ij}.

21. Equation (8-30) is obtained by substituting equation (8-29) when $m_2 = 0$ into equation (8-27).

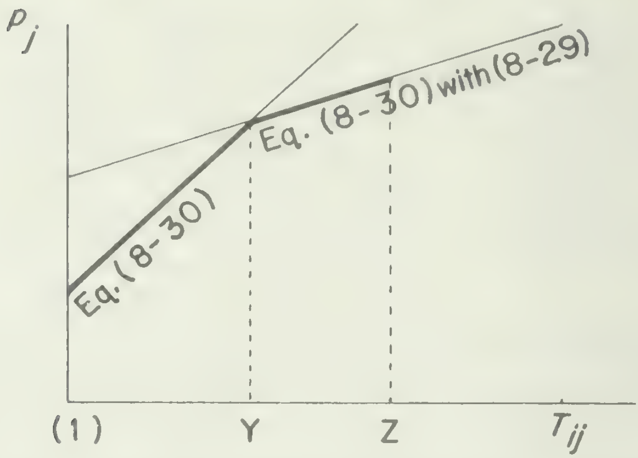


Fig. 8-11 The combined DPS of a firm at production center (1).

spatial competition, each linear segment of which relates to a constant level of competition.

Of course, the number of firms does not accurately measure competitive impacts, and different interpretations of the meaning of m_1 and m_2 can easily be presented, with modified results shown accordingly. Moreover, given a variation in competitive impacts, the DPS in fact becomes a combination of the schedules relevant when m_1 and m_2 are taken as constants along with m.

Two possible variations of competition are shown in Fig. 8-12. In panel (a) a varying number of sellers based at site 1 effect sales to market points T_{1j}; indeed, their competitive advantage is quickly dissipated as distance cost from this plant site rises above zero; in panel (b) the sellers are assumed to be subject to comparatively little competition from distant rivals over a larger portion of the market space.

The soporific effect of exhaustive categorization of possibilities often outweighs desirability of detail. But one remaining situation nevertheless warrants recording here. We refer to the possibility that one production center may be located off the transport line connecting the rival production center with a set of buyers. Though this is an elementary observation, it is not a simple statement to observe that the DPS to buyers along a transport route may assume many slopes when some competitors are located elsewhere than along that route. It may curve upward, or downward then upward, or be linear *when viewed in the perspectives of the sellers located on that line.* Exact answer to curvature depends on where the rival center is located. Manifestly, a myriad of locational competition possibilities exist and many more would warrant evaluation if a complete picture were needed. Present requirements are,

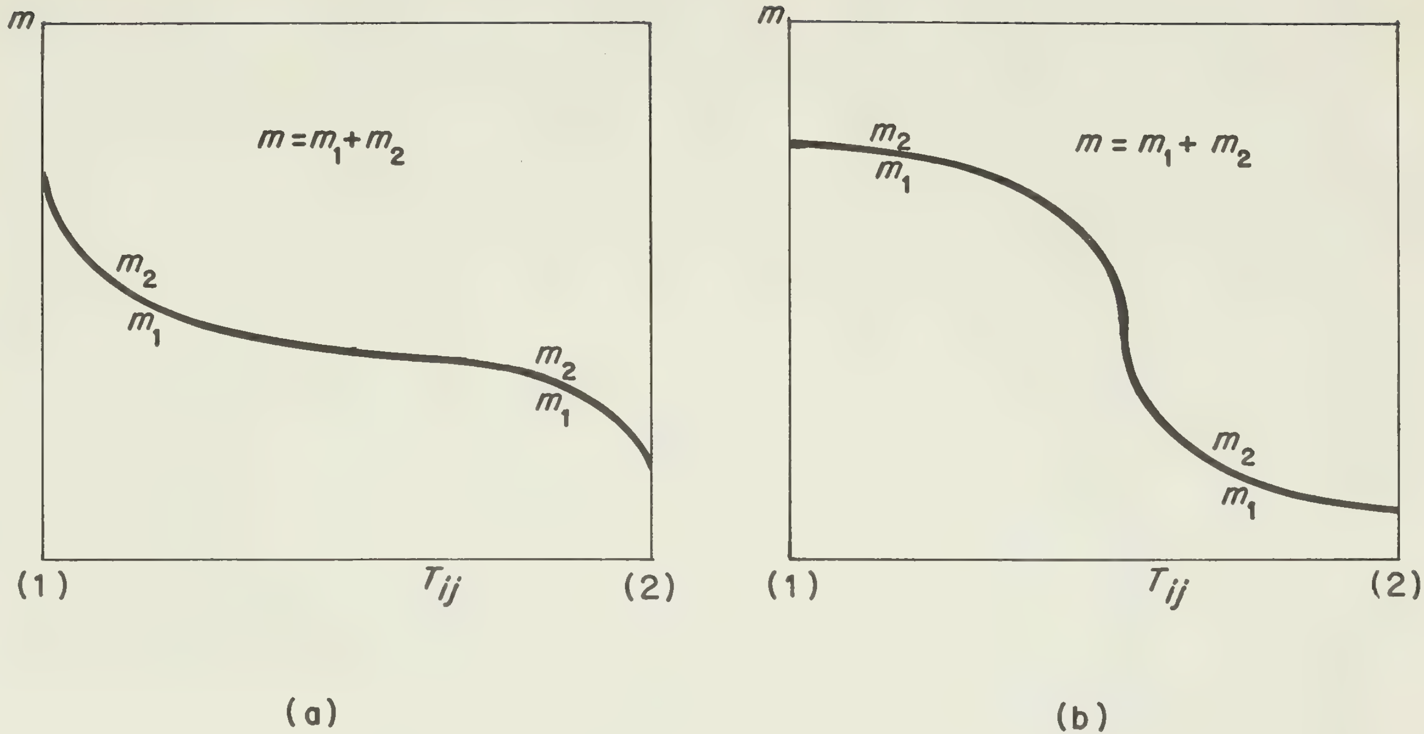


Fig. 8-12 The competitive variations.

however, fulfilled without such detail, and there remains a need only for a brief, concluding statement.

III

F.o.b. pricing results when competition at the production site is intense and there are few, if any, rivals located at a distance or distances from this center. But uniform pricing results when alternative and equally competitive production centers dot the landscape, since average transport costs throughout the space will be constant. Analysis of delivered prices under conditions where competition varies with distance revealed a DPS which is a combination of schedules, provided the total force of competition is constant over the space. The delivered-price schedule actually curves downward at points where competition increases at greater distances from a given production point. Rather than reflect uniform pricing, as when competitors are located on either side of the buyers, the DPS will be concave from below.

The more general case of buyers at nonmonotonically varying distances from different production centers (i.e., in effect the case of sellers located over the plain rather than only along a line joining the competitive production points) is characterized by a DPS which curves upward, or downward and upward, etc., depending on the exact location of the production centers. Viewed from the standpoint of the firm(s) located on the buyers' distribution line, the DPS can be expected generally to curve upward with increases in T from both centers, while from the standpoint of the distant seller(s), it generally curves downward with respect to buyers located closest to the rival production center.

The diversity of spatial price relationships and patterns possible in the free enterprise space economy reflects the importance of competitive structures on spatial prices. Moreover, we know from x in (8-27) that different demand curves affect the DPS. Placing everything together can lead to price gyrations over space which would be very difficult to administer. An empirical study currently in process suggests (at the time this material is being written) that uniform delivered and f.o.b. mill price schedules typically characterize the spatial pricing of firms except where custom-made orders are possible. In that event, the effect of special forces (demand and competitive differences) can be hidden behind the assertion that "our prices are subject to individual negotiation."

It warrants final mention that the analysis of this appendix has been based on a parametrically fixed number of firms as well as given locations of firms over space. However, new prospective firms may enter the market and/or the established firms may relocate, seeking to obtain

profits or increase their profits. In the process, some large firms may locate together at a point which minimizes the transportation costs to all or at least to many major market points; at the same instance, some smaller firms may locate at a distance from this production center and sell to selected nearby market points. It is even possible that some very small firms will locate at a particular market point and sell only to this market. In the long run, excess profits are wiped out and all viable firms will have located most efficiently. Though Appendix II, which now follows, evaluates the welfare properties of the space economy essentially in terms of classical marginal propositions, it will be seen that the locational efficiency suggested here is counterpart to the technological efficiency that we shall next propose.

Appendix II: *Plant location theory and spatial prices*

I. *Introduction*

We observed in Chapter 1 that regional economics consists of many subjects, one of which is location of industry. It was, in fact, suggested there that the inquiry of this book comes very close to the subject of plant location, since the firm's pricing patterns over economic space influence its location practices.[22] In this same context, the chapter just concluded implicitly assumed a limitless number of locational alternatives at a distance, an assumption which enabled us to conceive of locations as "given" while focusing attention on the price and distance variables. This particular framework of thought generates a question. Would the location theory to be set forth in this appendix, and would the "given" locations assumed previously and again in Chapters 9 and 10, be altered (or in fact be fundamentally different) if economic space were actually dotted by *competitive* oligopolists who price discriminatorily rather than f.o.b. mill? Though this particular question cannot be evaluated directly herein, much less in detail, the underlying effects will be suggested from time to time and otherwise apparent to the reader. Moreover, we shall break somewhat in Chapter 11 from the practice of assuming fixed locations, as we move a little bit forward there towards the goals of (a) determining the impact of discriminatory pricing on locations, and hence (b) its impact on the "given" locations assumed elsewhere in this book.

22. As has been shown elsewhere by one of the authors of this book, collusive oligopolistic pricing practices, such as the basing-point system, induce a different location pattern than that of f.o.b. mill pricing. See Greenhut [9, chap. 2].

To set the stage for the particular inquiry of this appendix, as well as to provide background thoughts for Chapters 9 and 10, we shall utilize parts of a paper on location theory published years ago by one of the authors [10]. At the risk of unnecessary repetition, let us stress that the passages to be included were written originally under the assumption of f.o.b. mill pricing (i.e., model II). Notwithstanding this background, the reader should keep model I in his mind (i.e., our spatial price discrimination model) as he reads the following extracts. In fact, it is suggested that he can conceive of model I as if *it had been the pricing system assumed in the materials that follow.*

II. *The theory of locational interdependence*

The increasing awareness in recent years of the limitations of the von Thünen–Weber purely competitive spatial framework led to a new approach in the early part of the twentieth century to the problem of plant location. Under the influence of Fetter [5], Hotelling [13], Lerner and Singer [15], Smithies [25], Chamberlin [2, app. C], and other writers, interest centered upon locational interdependence. This analysis was predicated upon the theory of duopoly. It abstracted from cost[23] and explained the location of firms as the endeavor to control the largest market area. The methodology and conclusions were in contrast with Weberian findings, but the problem was essentially the same.

In the Weberian framework, price is given and buyers are assumed to be concentrated at point-formed consuming centers. Because no sales-price advantages are available at locations closest to the selected market, the location decision in reference to a given market center rests solely upon cost. In the Fetter-Hotelling framework, delivered price varies with location, for buyers are assumed to be scattered over an area. Each firm (subject to identical costs and following the same price system) is able to sell to proximate buyers at delivered prices which are lower than those of their rivals. This industrial pattern signifies exclusion of competitors from the market area surrounding a seller's plant. The location of each seller thus depends on the number of consumers he can monopolize at the different locations. The attempt to control that number of buyers (market area) at prices which yield greatest returns is the driving force behind orientations to market areas. Under this framework, the size of the market area belonging to a firm is determined by the locational interdependence of the firms.[24]

23. These theories assumed that the cost of procuring the raw materials and processing them were equal at all locations.

24. Each seller becomes a locational monopolist when consumers are scattered around a seller and he prices f.o.b. mill (or for that matter discriminatorily). In Weber's system, sellers compete actively for the same market, as in perfect competition.

The assumption of a market area leads the writers on locational interdependence to an appraisal of the influence of (i) the shape of the industrial demand curve, (ii) the shape of the marginal cost curves, and (iii) the height of the freight rate on the location of industry. These factors influence the conjectural hypotheses of entrepreneurs about the location policies of rivals; they, along with others, such as the history of the locational competitiveness of an industry, determine the degree of dispersion when conditions of sale are perfectly competitive except for the impact of space. The effects on theory of these locational forces are shown in the following models.

Assume:

1. Buyers are uniformly distributed over a linear market. (The conception of buyers distributed along a line instead of over an area simplifies the problem without affecting the nature of the conclusions.)
2. Buyers are indifferent to sellers; neither sellers nor products are differentiated, except for location.
3. Sellers are indifferent to buyers; all buyers are homogeneous in every respect except location.
4. The seller's procurement and production costs are everywhere equal. Marginal costs for competitors are constant. With no great loss in realism, average and marginal costs are assumed for the present to be not only equal but zero.
5. Each competitor sells on a nondiscriminatory f.o.b. mill basis; that is, the seller sets the same mill price on his sales to all buyers.
6. Each competitor is capable of supplying the entire market.
7. Cost of transport is at the same rate per unit of distance throughout the market area.
8. Each competitor is free to move his location instantaneously and without cost.
9. The security motive does not influence the locational choice.[25]

(i) If the demand for the product is *infinitely inelastic* and each buyer purchases one unit of product at one unit of time, duopolists will concentrate together at the center of the market area. From this vantage point each can provide the remotest extremities of the market while limiting the rival to no more than his proportionate share of the market.[26] On the other hand, an *infinitely elastic* demand disperses sellers of very

25. It is readily understandable that the security motive may determine the location. For a discussion of this motive, see Rothschild [22] and Tintner [27]. Of course, Machlup [18, pp. 51–56] suggests that the security motive is part and parcel of maximum profits. But distinction between the two does have practical as well as theoretical basis. See cases 5 and 6, Greenhut [7].

26. This is the celebrated Hotelling case [13, p. 52]. From any firm's standpoint, the sharper competition with rivals due to proximity is offset by the greater number of buyers over whom it has an advantage. As we will see later on, the *general* conclusion for this case (when there are more than two firms) is dispersion.

limited output capacity to the point where they locate in perfect harmony with the location of buyers.[27]

If the demand curve for the product of an industry is assumed to be of the negative-sloping straight-line variety, firms will not localize at the central locations, nor will there be any great assurance of their perfect dispersion. Only in the long run would they tend definitely to locate like the multiplant monopolist, who, if he has two plants, locates one at the first quartile and the other at the third quartile of the linear market, and if he has three plants locates one of them at the first sextile, one at the third, and one at the fifth, etc. This scattering minimizes freight costs and maximizes profits. Its necessary precondition is the entrepreneurial conjecture that any attempted incursion in the market area, either by way of lower price(s) or locational movement to the center, will be met by the other competitor. This means that rivals must have knowledge that the quartile, sextile, etc., type of location is most profitable for each firm. In turn, this understanding is determined by several factors, among which the shape (straight-line, concave, convex) of the demand curve takes a leading role.

Assume in this context that the same maximum price applies to all demand curves regardless of demand-curve type (i.e., straight-line, concave, convex), and further assume that the demand curves are tangent at the spaceless equilibrium profit-maximizing price. Demand-curve convexity may then be seen to add an element of doubt concerning the rival seller's conjecture of which location would maximize his profits. Convexity from above signifies that absorption of freight costs in mill price will be great and that sales to the peripheries of the market area *may be* obtainable from locations central to the entire area. It is thus doubtful whether a rival would actually move to quartile locations even though the result would be advantageous. *(May we insert the thought that discriminatory pricing would further stimulate such doubt.)* Conversely, when a demand curve is concave from above and is finitely limited to the same maximum price as the linear or convex demand curves, mill price will be greater, on account of freight, than in the other cases, and *may even be greater* than the mill price that would have prevailed in the absence of freight.[28] Such a pricing result promotes the quartile, sextile,

27. See Morris A. Copeland [4, pp. 3, 23], who assumes an infinitely elastic demand curve and concludes that sellers will disperse in accordance with the dispersion of buyers. Of course, a one-to-one dispersion would require the assumption that each buyer can purchase the entire output a seller is able to produce at the relevant marginal cost price. And see A. P. Lerner and H. W. Singer [15, pp. 145–86], who assume an inelastic demand over a range extending from zero to a finite upper limit. And "on the assumption that A takes B's position as given" (i.e., location and price), they conclude that A will not behave as in Hotelling's model, but will remain at a distance or attempt to drive the rival out of the market (p. 154).

28. This result can be understood by realizing that the more concave the demand curve, the farther to the left will be the intersection of marginal revenue with a marginal cost

etc., type of locations, since any given firm may expect a competitor to seek distant sites in order to gain sales at peripheral points of the market area.

(ii) The shape of the marginal cost curves could be a similar determinant of locational interdependence (concentration or dispersion) when buyers are assumed to be scattered over an area rather than concentrated at a point. If marginal costs increase with greater shipment to any given point of the market area, freight absorption in pricing is great. It follows that the impact of freight cost on price may be so slight compared to the price(s) that would exist without spatial effects that firms at a central location may well be able to supply the peripheries of the market. To the extent that sellers seek large sales radii, doubt therefore exists whether a rival would move to distant sites. Conversely, when from a long-run viewpoint marginal cost curves are declining, the freight absorption that ultimately takes place in pricing over space is slight; in fact, the underlying mill price may actually increase because of the freight cost.[29] It follows that there is greater assurance in this case that a rival will move away from the center of the market.

(iii) The height of the freight rate is a locational determinant which influences the cluster or dispersion of industry. Low freight rates (or any lowering of freight rates) leads to localization in the center, since the hinterlands are in this case more serviceable from a distance. On the other hand, high transit cost (or an increase therein) tends to disperse the affected industry.[30]

There has been dispute in the literature concerning location of three or more firms under inelastic demand. Hotelling [13] offered the view that sellers would localize at the center of the market. But this view has only slight merit under game theory; and while it could happen that firms would concentrate,[31] probabilities suggest some immediate disper-

---

curve that has been increased because of freight costs. The delivered price that is shown by this intersection can be higher than the mill price (which did not include freight cost in its consideration) by more than the freight burden that exists.

29. If a linear marginal cost curve is negatively sloping at the same rate as the average revenue curve, then the intersection between a marginal cost curve of this type, which furthermore has freight cost added to it, and the marginal revenue curve that is associated with the defined average revenue curve will leave a delivered price to the subject market point that is greater *by the amount of the freight* than the total price that would have existed in the absence of freight costs [9, pp. 305, 306]. There is no freight absorption at all in this case, and the mill price with and without freight cost considerations is the same. On the other hand, if the marginal cost curve is sloping negatively at a steeper rate, or at a more moderate rate than average revenue, the mill price that includes consideration of freight costs will be greater or less than the nonspatially formed mill price, as the case may be.

30. And see Schneider [23, p. 80], who draw the further conclusion that *ceteris paribus* "die Grosse des Absatzgebietes . . . dem Quadrate des Frachtsatzes umgekehrt proportional ist."

31. If plans must be made well in advance of any construction, so that land options and contracts to purchase materials, to float securities, and to borrow funds, when once ar-

sion.[32] According to Lerner and Singer, the dispersal influences which exist are formed by the need of each firm not to be caught between two others, that is, unless the distances between firms are either substantial or its opposite, microscopically small [15, pp. 176–79]. To these writers, the locations are unstable over time. The cycle may find three firms initially at the center, then two at one quartile and one at the other, then all three back in the center, etc. The underlying force behind the wave of relocations is the assumption by each competitor that the site selections of the others are fixed. Chamberlin [2, app. C], on the other hand, deduces the possibility that two firms may be located at opposite quartile locations and the third competitor at some place between them. He offers the opinion that the continual shifting of sellers seeking advantage carries the dispersion at least this far and perhaps farther.

For cases of more than three firms, a clustering form of dispersion as in the Lerner-Singer adoptions may arise, or else a more complete dispersion in the Chamberlinian tradition appears likely.[33] It is significant

---

ranged, can be ill withdrawn, the location problem under uncertain conditions becomes like most games; it involves hidden play. Only at a given moment does revelation of decision come about; and then, the revelation by one seller is followed closely by that of another. This high simultaneity in announcement of decision suggests that in the absence of collusion, rivals had to conjecture and plan without precise knowledge of each other's ideas. The uncertain conjectures of a firm form the basis for the theory of games and plant location. Consider in this regard the following.

The conjectures of any one firm must be that its two other rivals, operating under the same economic climate and subject to comparable expectations, will each plan to locate at a substantial distance from the other or else intend to locate close together. The numerical values to be assigned to these possibilities must exceed those accorded to the likelihood of opposite plans by each firm. Considering then only the cases of comparable plans, it may be that each firm with plans for dispersion will locate at the opposite quartile, or else of course their plans may not have succeeded and both may locate accidentally at the first or third quartile together. On the other hand, similarity in plans may suggest intent to locate together; in this event, planned location at the market center is most probable. It is manifest that only in the last case would the third firm do well to select a site distant from the market center. In the other cases, central location suggests advantage. Because there are more alternatives in favor of central location, the minimax site becomes this one. Of course, any business man may then entertain the thought that if the other two act according to this minimax site, its own maximin site would become the distant location rather than the central one; but comparably, this same extra thought may be conceptualized by the others and hence would be discarded. Further, let it be noted that because accidental concentration of all three at the first or third quartile would encourage numerical expansion of the industry, the firms would tend to locate together near the center according to this line of thinking rather than at the quartile type of site.

32. Any tendency to localize fully at the market center pursuant to the game conjectures indicated in note 31 above is weakened by the fear that the firm might be hemmed in between the other two firms. Thus any one firm will plan location near the center, but somewhat to the right or left of center. Accidentally, the Hotelling solution of full concentration may arise. But chances are that two firms will be to the right (or left) of center, with the other firm sited to the left (or right) of center. Indeed, if the ends of the line are joined to form a circle, probable separation of all firms is sharply revealed. And see Chamberlin [3, pp. 19–20].

33. Over time, on a circular street or over a circular market area there exists a stronger tendency to full dispersion. If the fourth, fifth, and still later firms locate after leaders are

that the tendency toward separation becomes stronger as the number of firms grows larger. It always remains possible, however, that the great fear of being hemmed in may stimulate tacit agreement to localize together at the market center in sharply concentrated form. Obviously, such extreme localization would be somewhat unusual as well as unstable over time.

Once the Hotelling assumption of an infinitely inelastic demand curve is dropped, it becomes apparent that firms will disperse in a more ideal way. Such locations follow the limitation set by a finite demand. Indeed, if in addition to demand elasticity, the cost of transport and the marginal costs of production are assumed to be high, the scattering of firms will proceed quickly in greater conformity to the location of buyers. Each firm would be more willing to announce its plans for distant location without waiting to play the role of cat or mouse.

While economists generally have postulated identical production costs everywhere, the above analysis can be broadened to include the case of divergent production costs at alternative sites. (This type of model may be considered with greater advantage in a separate section, Section III below.)

To summarize and appraise the theory of locational interdependence, this theory has pointed to the following main conclusions. (1) The tendency to disperse depends upon the height of the freight cost, the elasticity of the demand function, and the characteristics (slopes) of the marginal costs. These factors, along with historical practice, determine the degree of competition in location. (2) Each seller seeks to control the largest (most profitable) market area as the actual location selected is determined by the type of interdependence existing between the seller and his rivals. (3) Each seller becomes a spatial monopolist (to some extent at least) when sellers and buyers are separated geographically from rivals and the cost of distance is significant. (4) Effective demand varies at alternative sites because of freight costs and the *location* of rivals. (5) Locational interdependence theory is designed mainly for manufacturing-plant locations where raw materials and labor supply are somewhat ubiquitous and where market share is the vital variable factor within the sphere of influence of the firm. The overall theory probably also applies to locations of functional middlemen and to certain special wholesaling and retailing functions. It can even be introspected that the individual who derives his satisfactions (religious, social, climatic—including amount of space and air—entertainment, etc.) over an area is in a position similar to that of the firm. As with the owner of a business establishment, he plays for the right location within a landscape of diverse utilities. (6) Three or more firms locate in much the same way as two. All

already in the market, the dispersion may be complete or in clustering form, depending upon the pattern that prevailed prior to the entrance of the later firms.

other things being equal,[34] the force of dispersion is more pronounced the greater the number of firms.

The shortcoming of the theory of locational interdependence is patent: by design it makes an abstraction from cost that is similar to the abstraction from demand of the least-cost location theory. It yields, therefore, only a one-sided theory which reveals a special type of explanation of underlying forces of location in economic space. The following section attempts to avoid the specialization in assumptions of Section II. It treats cost and demand as variable factors that govern the selection of sites. This examination uncovers the intense uncertainty that develops over rival strategies.[35]

III. *The theory of maximum-profit plant location*

Suppose the B location of Fig. 8-13(a) – at the center of the market – is the least-cost position and that other sites situated away from B entail heavier procurement and production charges. Where will firm A locate?

If the costs of production and procurement at locations other than B exceed transportation savings from these sites to any buyer, no location other than B is feasible. Fig. 8-13(a) reveals this situation (the BP and AP′ stem include equal mark-ups above cost). As a general rule, it follows that where P_B (price at B) + T_B (transportation from B) is $< P_A + T_A$ at any location other than B, all firms must locate at B. *(Results would be essentially the same if the slopes or shapes of delivered price curves were changed.)*

Where the procurement and production costs are significantly larger at other locations, but an untapped distant segment of the market remains which can be supplied most economically by a plant situated near to this section of the market, a location at A in Fig. 8-13(b) is feasible. In this case, the firm may charge a higher mark-up above cost so as to ex-

34. Example of inclusion in the *ceteris paribus* assumption is the type of product being sold, shopping good or convenience good.

35. The line problem of location has one other aspect besides its emphasis on locational interdependence. This side comes into play under the model of many many firms in the market. In this case, each new participant is forced to locate in accordance with the remaining availability of sites. In turn, the actual number of firms in the market and their dispersion become the central problems for analysis.

Resolution of the number of firms and their dispersion in the market depends essentially over time on the factors of convenience in buying and the extent to which economies of scale prevail. These underlying determinants (demand and cost curves) are shown by Lewis to lead toward the ideal number of firms in the market and associatively toward the ideal size of firms when economies of scale are tied up with a quite elastic demand. This tendency, he says, is lessened when scale economies are related to a less elastic demand. Space relations are thus shown by Lewis to be one aspect of monopolistic competition, and the tangency solution which he drew involves higher than optimum cost. See Lewis [16].

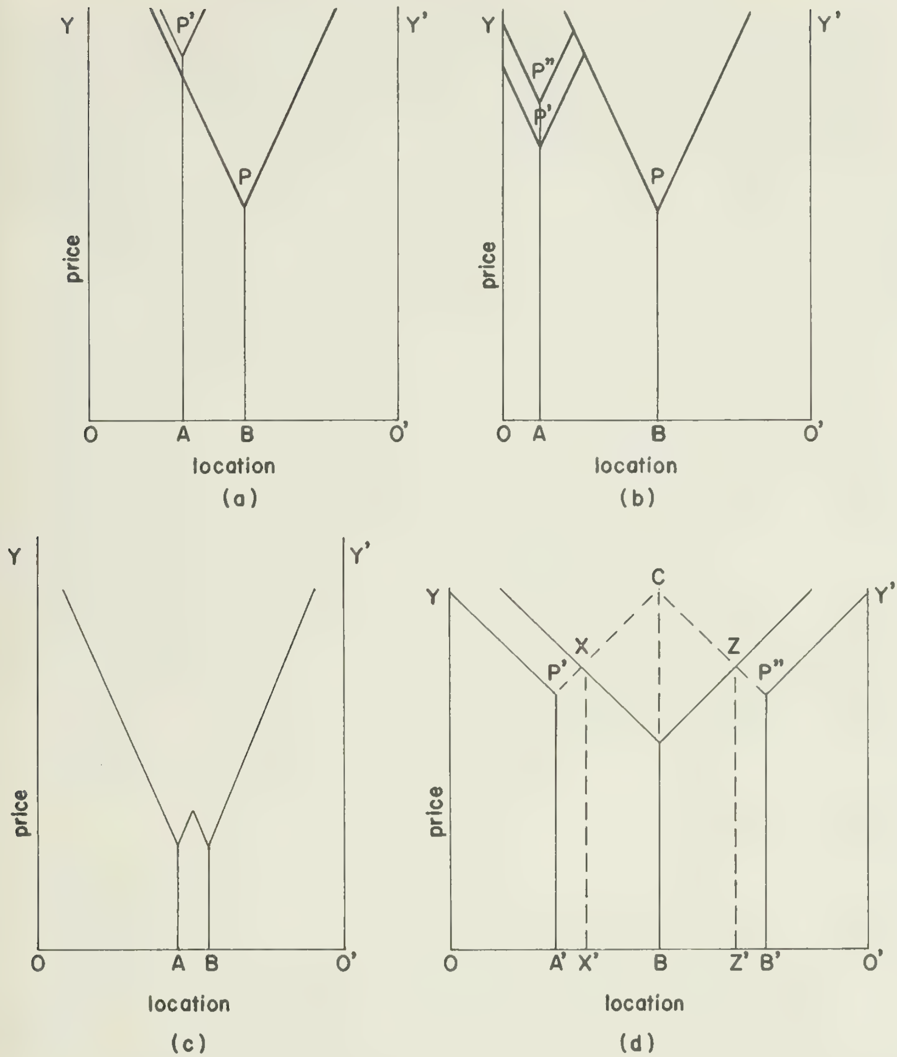


Fig. 8-13 Feasible location under alternative cost conditions.

tend the AP′ stem to AP″. High-cost locations made in reference to such market segments would generally be adopted by small-sized plants.[36] We might add the thought here that the distant location A *must* possess a cost advantage with respect to B in order to enable its firm to win a price war; and see above, Chapter 7, Section III. On the other hand, where

36. Ritschl [21, pp. 813, 829, 870, 848 . . .]. Faraway sellers can exist even though they have higher production costs and a smaller sales radius. The freight rate protects these firms.

owners plan larger-scale enterprises, they must compete actively at or near B in order to attain their size objective; see Fig. 8-13(c).

Under the assumption of equal costs, it has already been noted, quartile positions result initially or in the long run when each of two firms (owned by A and B) assumes that any price set or any location adopted by the one will be met in identical pattern by its rival. However, if costs are not identical at all locations, this quartile movement does not necessarily take place. The determining factors are the sales, the mill prices, the costs at the two locations,[37] and the conjectural hypotheses of the locators. Significantly, imperfect knowledge by A and B of each others' costs arises when costs are not uniformly equal over space. This combination introduces a special element that must be considered in theory.

The essential difference between the cases of equal and unequal cost is the element of extra doubt. Though past performances of B may indicate quasi-monopolistic policies on his part, the future may reveal antithetical inclinations in him. Quasi-monopolistic practices are based upon economic advantages gained by nonactively competing firms. But B may embrace competition in location in this market *if* convinced that he cannot lose thereby. Such a conviction follows the belief that the high costs at the quartile or other locations cannot be made up by savings in freight, the consequently smaller freight absorption, an anticipated lower average delivered price, and the expected increase in sales. If this is the conjecture of B, he has nothing to gain and everything to lose by location at a site other than the central (least-cost) site. Conversely, A goes directly to the quartile or other positions only if convinced that B always locates symmetrically.[38]

The case of unequal costs suggests two new principles. (1) Quasi-monopolistic tendencies in price are not so closely associated with like practices in location when costs are unequal everywhere. The less definite relation results from the conjectural hypothesis discussed above: the uncertainty of the situation. To this must be added a new consideration, which is also the second principle. (2) The rivalry from substitutable products is probably more pronounced, more influential, and therefore a determinant of location when costs are dissimilar at alternative locations.

The assumption that there is a least-cost location with respect to the

37. Sales at the quartile location are greater if the delivered price to a large number of buyers is sufficiently lower from this location than the delivered price from the least-cost site. The comparison between delivered prices depends upon the extent to which savings in freight compensate the higher marginal costs. Finally, the mill price and associated sales must be so much larger as to raise the net average revenue curve sufficiently in respect to the higher average cost curve. An increase in sales alone does not make it profitable for firms to locate at sites of higher cost.

38. The heroic assumption of zero costs in relocation would be necessary to justify a short-run quartile location.

entire market area suggests the possibility that similar industries are having similar experiences. If it can be presupposed that the least-cost location for the industry in question is also the least-cost location for many other industries, and particularly for an industry producing similar goods, the location of substitutable (heterogeneous) products becomes increasingly important.

This determinant of location can be seen readily by reference to Fig. 8-13(d). In that figure, quartile locations cause an increase in price between X′ and Z′ compared to the price which would exist if a central location were adopted. The increase in delivered price may cause extra losses in the district between X′ and Z′ because buyers shift to the substitutable product. Not only does movement away from the central location (e.g., to A′ and B′ in Fig. 8-13(d) lead to the normal losses between X′ and Z′ on account of higher delivered prices, but extra losses result from the high cross-elasticity of demand at the central site.[39] Furthermore, there is a smaller probability that a firm (industry) producing a substitutable product will locate symmetrically (disperse) with rivals, because the company has formed a custom by differentiating its product. These conditions mean that there is less likelihood that heterogeneous firms will pursue quasi-monopolistic location policies in their dealings with each other. Manifestly, where firms are large-scale producers and their costs vary widely among locations, they will concentrate, *ceteris paribus*, in the particular city or district which offers lowest average cost for the total output planned by the owner.[40] The location of such a firm, i.e., a large firm, is accordingly made with particular emphasis on the least-cost site in relation to the whole market area. Such firms do well in emphasizing the location of their rivals. Only the foolhardy dare chance the initial move if doubt exists as to the probability of symmetrical locations.

The location of the smaller firm is somewhat more flexible according to this reasoning. These firms seek smaller market segments and therefore may locate at the least-cost position in reference to a given particular segment of the entire market. The smaller firm chooses its site with greater freedom in this respect; it is more likely to avoid locations near

39. Respective locations at A′ and B′ raise the selling price between X′ and Z′ above that which would exist if the optimum cost location at B was chosen. The more effective utilization of the X′O and Z′O′ districts, *ceteris paribus*, may or may not compensate the loss between X′ and Z′, depending upon which segments of the market had the higher cross-elasticities. To the extent that point B represents a highly industrialized location, the loss to rival industries may be greater than the hinterland gains. On the other hand, where costs are equal everywhere, the problem of substitutable products is most pronounced at the market extremities. This latter result is derived from the fact that delivered prices are highest at places most distant from the seller's factory, and therefore the demand is, generally, most elastic at such sites. Equal costs indicate quartile locations for homogeneous firms; unequal costs may not suggest the same result.

40. See Florence [6, p. 84]. The large firms within an industry concentrate at given points.

competitors; a priori, smaller plants generally are found in the less industrialized areas.[41]

The inclusion of cost and demand factors in one model points out the need for a broader statement of the determinants of plant location than one which concludes that firms seek the least-cost location, or one which holds that firms seek the location offering the largest market area. This need is fulfilled by the concept of the maximum-profit location.[42] By definition, this location relates to that site from which a given number of buyers (whose purchases are required for the greatest possible profits) can be served at the lowest total cost. And while the lowest level of average production cost at this site may be higher than that which exists at alternative ones, the monopolistic control gained over larger numbers of buyers (spread over a market area) makes it the maximum-profit location at the optimum output.[43]

This definition of the locator's objective recognizes the fact that one location may offer lower manufacturing unit costs at a given output than another, but that the relative positions may be reversed as the market area of the firm expands and differentials in transport cost appear. For example, high-cost manufacturers who extend their operations by water transportation may gain lower total unit costs to distant buyers than competitors who are linearly nearer to these buyers but are forced to use the railway or highway networks.

This concept of the maximum-profit location does not have to be confined to an analysis in which the demand or cost factor is held constant. It lends itself readily by definition to an examination of cost and of locational interdependence. When uncertainty prevails, implicit modification of the concept is, however, required. A minimax force arises which finds impropriety in the more daring movement to the quartile. As an implicit part of the concept of the maximum-profit location, the minimax principle finds effect in its tendency to draw firms together. The strategy of games may, in other words, suggest concentration when dispersion otherwise would result. Maximum-profit location theory appears as a consequence to be adaptable not only to manufacturing-plant locations but also to locations on other levels, including the retailing level as well as that of the individual.

It is manifest that the location theory just described offers insight into

41. This statement, of course, excludes from consideration the service companies. See Florence [6, p. 84] for statistical verification of this deduction. Small plants disperse more than the large firms. They locate generally in the less industrialized areas.

42. This maximum-profit location concept contains many sides. Essentially, it excludes psychic income factors.

43. Weber's least-cost location fits the above definition in part; its compliance is limited, however, because Weber's least-cost location must emphasize the cost of procuring, processing, and distributing goods to a *given* buying point and thus it really does not (nor can it) convey spatial implications.

an objective of plant location and the ways and means by which plant managers seek to achieve this objective in their selections of plant sites. It points to the following main economic conclusions. (1) When firms sell to a given buying point, they seek the least-cost location in reference to this consumption center and ignore the locations of rivals in selecting their plant sites. (2) When firms sell over a market area, their site selections are influenced greatly by the location of rivals. (3) In selecting a plant site, each firm seeks that place which offers the optimum sales output at a cost that cannot be matched elsewhere. (4) When firms sell over a market area, and unequal costs exist at alternative locations, the minimax force of concentration becomes an inherent part of the locational patterns. (5) When firms sell over a market area, the tendency to disperse depends upon the height of the freight cost, the elasticity of the demand function, the characteristics (slopes) of the marginal costs, the degree of competition in location, the degree of competition from substitutable products at the various locations, and the homogeneity or heterogeneity of the firms belonging to the industry. Of course, such market imperfections as time of delivery, personal contacts, custom, equalizing and other types of discriminatory price systems, influence the localization of industry and distort the findings listed above. These market imperfections tend mainly to yield special theories of location, not the general theory which is required for understanding the underlying trends.[44] At the same time, they suggest that different ethnic and religious groups might present some special unique factor(s) of consideration that would distort any economist's presentation of the location of individuals founded upon the pure basis of general cost and utility functions.

IV. *The effectiveness of the spatial distribution*

The private enterprise economy requires a maximum-profit type of location theory in which uncertainty (with its minimax principle) is a main variable. Only in this light are the underlying location forces of such an economic system revealed.[45] This framework, as outlined in Section III above, must be the cornerstone for a general theory of location in a free enterprise economy. It appears capable of explaining practically all business locations as well as practically all personal site selections. Indeed, the essential difference between locations of individuals and firms would seem to be that the individual's cost factors are formed mainly by

44. And see Greenhut [8] for an example of rejection of certain special locational data and subsequent derivation of a generalized theory.

45. Locational interdependence clearly involves a non-zero-sum game. And see Nash [20].

transportation and land rental charges while his demand (utility) factors are influenced by family and religious relationships and similar forces. The spatial regularities shown in demographic studies,[46] the associated regularities exampled in the investigations of business locations,[47] and the theoretical expectations of spatial regularities[48] suggest that all location rests on the same underlying principle. This principle finds roots in the cost and demand factors of the business firm and, their counterpart, the scarcity/utility factors which govern the individual's quest for maximum satisfactions. In very much the same way that Pareto's system of general sociology is a reflection of general economic theory, the location of individuals is a reflection of the location of firms. A general theory of plant location in a private enterprise economy may very well include the person, with the individual's quest for space, public facilities, neighborhood, and neighbor type creating a force of uncertainty reflective of uncertainties of business location. In any event, the theory of games has a spatial content: the theory of industrial location. To the extent that we possess an economy marked by a mixed land surface heterogeneous in part (e.g., as in Weber's framework of thought) and homogeneous elsewhere (e.g., as in Lösch's framework), to the extent that we experience inequalities in costs (Weber: other than those that arise only from agglomerating advantages) and are subject to some spatial scattering of consumers (Lösch: open area and full area sectors), a combination of theories would appear to be applicable. May we suggest finally that to the extent that the economic landscape is homogeneous in part and heterogeneous elsewhere, a rather noticeable intra-industry scattering and inter-industrial concentration may be expected. Most fundamental, the dispersion of firms and the agglomeration of firms tend to conform to the cost pattern and demand forces prevalent over the subject space.[49]

V. *Conclusion*

The material presented above was derived with very slight editorial additions from a paper written nineteen years ago; certainly, it is not the "last" word on the subject of location theory.[50] A significant portion of

46. Zipf [29], Stewart [26], and Singer [24].
47. Lösch [17] and Florence [6].
48. Tintner [28] and Isard [14].
49. A detailed presentation and elaboration of this thesis is given in [9]. And *suggestive* basis for our thesis is provided in Appendix II to Chapter 10, below.
50. A comprehensive updating of writings on location economics which tie production functions to uncertainty and location theory itself is in preparation at this time. But as was noted above, inclusiveness, precision in theory, etc., do not mark our present objectives, since we are simply proposing here the idea that what might be written about plant location in the backlight of the nondiscriminatory f.o.b. mill price system also applies to a price discriminatory system.

that paper dealing with long-run distributions and the effectiveness of the free enterprise economy was, in fact, omitted, since such discussion would have been tangential to the interests of this book. (See also [10, p. 9]; also along these lines the Appendix to Chapter 10, below. In particular, see [11, chaps. 12 and 13].) The *sole basis* for including the above materials was to suggest that the fundamental properties of plant location theory (whatever the most general or specific form of the theory may be at this time) would be largely unchanged by the prevalent use of the discriminatory spatial pricing model I rather than the classical f.o.b. spatial pricing model II. This is not, of course, to say that the number of competitive firms and the distances between these competitors would be unchanged by different spatial pricing systems. In fact, as mentioned at the outset to this appendix, we shall find in Chapter 11 that these relationships do change. We simply propose here that the fundamental "outlines" of the theory, i.e., the locational substitution among costs, between cost and demand, between these forces and the uncertainty in the market, etc., would be essentially unchanged as a result of the use by spatial competitors of model I rather than Model II. This simple thought is not equivalent, however, to simple theory, nor is it reflective of only an elementary effect on economic development and welfare. The simple thought is therefore not a complete one, but merely introductory to the subject of industrial price schedules *and* plant location.

References

[1] M. Beckmann. *Location Theory*. New York, 1968.

[2] E. H. Chamberlin. *The Theory of Monopolistic Competition*. Cambridge, Mass., 1946.

[3] E. H. Chamberlin. "The Product as an Economic Variable." *Quarterly Journal of Economics*, 1953, 67:1–29.

[4] M. A. Copeland. "Competing Products and Monopolistic Competition." *Quarterly Journal of Economics*, 1940/41, 55:1–35.

[5] F. A. Fetter. "The Economic Law of Market Areas." *Quarterly Journal of Economics*, 1924, 38:520–29.

[6] P. Sargent Florence. *Investment Location and Size of Plant*. National Institute of Economic and Social Research, Study VII. London, 1938.

[7] M. L. Greenhut. "Observation of Motives to Industry Location." *Southern Economic Journal*, 1951, 18:225–28.

[8] M. L. Greenhut. "A General Theory of Plant Location." *Metroeconomica*, 1955, 7:59–72.

[9] M. L. Greenhut. *Plant Location in Theory and in Practice*. Chapel Hill, N..C., 1956.

[10] M. L. Greenhut. "Games, Capitalism and General Location Theory." *Manchester School*, Jan. 1957.

[11] M. L. Greenhut. *A Theory of the Firm in Economic Space*, 2d printing. Austin, 1974.

[12] E. M. Hoover. "Spatial Price Discrimination." *Review of Economic Studies*, 1937, 4:182–91.

[13] H. Hotelling. "Stability in Competition." *Economic Journal*, 1929, 39:41–57.

[14] W. Isard. "A General Location Principle of an Optimum Space Economy." *Econometrica*, 1952, 20:406–30.

[15] A. P. Lerner and H. W. Singer. "Some Notes on Duopoly and Spatial Competition." *Journal of Political Economy*, 1937, 45:145–86.
[16] W. Arthur Lewis. "Competition in Retail Trade." *Economica,* 1945, 12:202–34.
[17] A. Lösch. *Die räumliche Ordnung der Wirtschaft.* Jena, 1944.
[18] F. Machlup. *The Economics of Seller's Competition.* Baltimore, 1952.
[19] E. S. Mills and M. R. Lav. "A Model of Market Areas with Free Entry." *Journal of Political Economy*, 1964, 72:278–88.
[20] J. F. Nash. "Non-Cooperative Games." *Annals of Mathematics,* 1951, 54:286–95.
[21] H. Ritschl. "Reine und historische Dynamik des Standortes der Erzeugungszweige," *Schmoller Jahrbuch für Gesetzgebung*, 1927, 51:813–70.
[22] K. W. Rothschild. "Price Theory and Oligopoly." *Economic Journal*, 1947, 58:299–320.
[23] E. Schneider. "Bemerkung zu einer Theorie der Raumwirtschaft." *Econometrica*, 1935, 3:79–101.
[24] H. W. Singer. "The 'Courbes des Populations': A Parallel to Pareto's Law." *Economic Journal,* 1936, 46:254–63.
[25] A Smithies, "Optimum Location in Spatial Competition." *Journal of Political Economy*, 1941, 49:423–39.
[26] J. Q. Stewart. "Potential of Population and Its Relationship to Marketing." In *Theory in Marketing*, edited by R. Cox and W. Anderson. Chicago, 1950. Pp. 19–40.
[27] G. Tintner. "The Theory of Choice Under Subjective Risk and Uncertainty." *Econometrica*, 1941, 4:298–304.
[28] G. Tintner. "Homogeneous Systems in Mathematical Economics." *Econometrica,* 1948, 16:273–94.
[29] G. K. Zipf. *Human Behavior and the Principles of Least Effort.* Cambridge, Mass., 1949.

Part 4. Spatial Configurations and Nondiscriminatory and Discriminatory Pricing: The Monopoly and Competitive Oligopoly Cases

9. Spatial configurations: a review of the literature

I. *Introduction*

Chapter 2 proposed that the profit-maximizing f.o.b. mill prices of the firm in economic space change as the distribution, and number, of potential customers change. In particular, at any given linear distance from the seller's plant, the profit-maximizing f.o.b. mill price of the spatial monopolist is lower in the case of sales over a plain market area than when sales extend only over a line market (e.g., along a street). The structures are fully conformable in other respects.

Chapter 2 further proposed that given the distribution of buyers (and transport lines), the monopoly f.o.b. mill price system is substantially similar to simple (nonspatial) monopoly pricing. From Chapter 3 onward, the possibility of alternative buyer distributions over space was discounted via the assumption that all buyers were located along a line. This simplifying assumption can be continued no longer. Instead, full inquiry must be made into the matter of spatial pricing practices and their interrelationships with alternative buyer distributions. The present initial chapter on the subject, following in part along the lines of [2], will review salient highlights of theories of market sizes and shapes under conditions of nondiscriminatory pricing. Chapters 10 and 11 will develop the subject further, in particular by distinguishing ultimately between impacts on market sizes and shapes of discriminatory vis-à-vis nondiscriminatory pricing over economic space.

II. *A review of the literature*

A simple market area framework

The size of a firm's market area suggests the wisdom with which its plant location was selected. Thus, if (1) buyers are assumed to be evenly scattered over a plain and to have identical demand schedules, (2) the market is monopolistically competitive *because* firms are geographically

dispersed, (3) all firms charge nondiscriminatory mill prices which are respectively marked up above their costs by the same sum, and (4) freight rates on the final product are the same for all sellers, then the least-cost producer will quote the lowest mill price and hence will capture the largest sales territory, i.e., market area [3].

The theory of August Lösch

From the above basic beginning points, August Lösch proceeded to postulate a broad homogeneous plain with uniform transport features in all directions and with an even scatter of sufficient raw materials. Furthermore, he assumed that the agricultural population was distributed evenly over the plain, with individuals having identical tastes, preferences, technical knowledge, and production opportunities. The result of these assumptions, in his view, was the dotting of the plain with completely self-sufficient homesteads.

Lösch's analysis actually began with an inquiry into the situation in which a farmer considers supplying beer to others besides himself. Though limited by freight costs, this would-be entrepreneur has the advantages of larger-scale operations and specialization. If the market demand is sufficient, the farmer advantageously sells over a circular area; if the demand is too small relative to costs, this would-be entrepreneur fails to survive the short run or in fact may never have entered the market.

Under the assumptions that the market is imperfectly competitive *because* firms are geographically scattered, and that the demand curve for the individual farmers exceeds his costs at some levels of output, production is immediately profitable. But in time the entrance of competitors compresses the circular area into smaller and smaller size. This decrease in sales reflects the leftward shifting of the demand curve, which continues according to Lösch until the Chamberlinian tangency point is reached.

But, the circular market areas of firms cannot remain; even though these ideal shapes are close enough to touch each other, perfect merging is impossible. Accordingly, profits continue and new firms enter. It is only when the circular area is reshaped in the form of a hexagon[1] that the equilibrium conditions can be satisfied.[2] The ultimate balance in

1. The fundamental thought which led Lösch to this conclusion is that the regular hexagon has a basic advantage over the circle and other polygons. It is capable of using up the entire economic space without deviating from the ideal circular shape as much as either the square or triangle [6, p. 73].

2. Lösch's equilibrium is determined by a system of equations for which the first condition (1) is that each producer maximizes his gains. This requirement involves equality between marginal revenue and marginal cost. The next three conditions require the number of in-

spatial competition is realized only when the hexagon is so reduced in size that profits are completely eliminated.

III. *The profit-maximizing size of market areas: a simple model of the f.o.b. mill price system*

The contribution of Lösch

Lösch's contribution extends beyond the spatial equilibrium which he proposed. His work opened the door to the possibility of determining the exact sizes of market areas under conditions of spatial competition as well as spatial monopoly. The present section of this chapter goes forward in this direction.

The contribution of Mills and Lav

Though Lösch's argument stemming from the assumptions given above appears to be a telling one, E. Mills and M. Lav [7] objected to it. They claimed that the circle may prevail as the equilibrium market area shape even under conditions of competition over economic space. They presented, accordingly, a proof applicable to certain cost and demand conditions which implied that a circular market of a given size (in fact any regular polygon with more than six sides) would provide greater profits under competition than would a space-filling hexagon of the same size. In other words, the former type of market area, not the latter, could prevail. Their system of thought as well as that of Lösch may be evaluated via a simplified model of the f.o.b. mill pricing system which employs basically the same assumptions as those given by Lösch and which also conforms closely to the system presented by Mills and Lav. The assumptions warranting stress are these: (1) buyers are evenly distributed over a plain; (2) each buyer has the same linear demand curve with unitary slope; (3) marginal cost of production is zero; and (4) transportation cost per unit per mile is of unit rate. Via these assumptions, Lösch derived the hexagon as the market area shape which would result from competition in economic space, while Mills and Lav derived the hexagon

---

dependent existences to be maximized. (2) All areas must be served by at least one firm. (3) All extraordinary profits must disappear, because under the competitive free-entry assumption new rivals eventually will eliminate all rent-like incomes. And (4), pursuant to the zero-profit condition, the area served by each individual must be the smallest possible. The fifth condition (5) of equilibrium requires that any consumer on a boundary line be indifferent to the possible sources from which he can obtain a given commodity at minimum cost. Without this condition, the boundaries would overlap, and the points of indifference would take on zonal qualities. See Hoover [5, pp. 58–59] besides Lösch [6].

only after competition had been carried on to a relatively significant extent.

To buttress their claim, Mills and Lav noted that the demand function at any point in a plain may be specified under assumptions (1), (2), and (4) as

$$q = b - (m + K), \tag{9-1}$$

where q and $(m + k)$ stand for the quantity demanded by a given buyer and the delivered price respectively, the latter in turn being defined as the sum of the mill price m and freight cost per unit over any distance K, where distance K is expressed in terms of dollars.

Given any regular polygonal market of some size K_* (defined below) and given price m, total sales Q of a firm will be

$$Q = 2\rho \int_0^{\pi/\rho} \int_0^{K_*/\cos\theta} \{b - (m + K)\} K \, dK \, d\theta, \tag{9-2a}$$

where ρ stands for the number of sides of a regular polygon, K_* is the shortest distance from the seller in the center of the polygon to a boundary point, and θ is the angle in any triangle (formed within the polygon) which connects the seller's site to a given set of buyers.[(I)] This equation (9-2a) applies to any regular-polygon market shape.[3] For a circle market (not regular polygon) of size K_*, total sales Q of a firm will be

$$Q = \int_0^{2\pi} \int_0^{K_*} \{b - (m + K)\} K \, dK \, d\theta. \tag{9-2b}$$

In this simplified version of the Mills-Lav derivation from Lösch, total profit net of fixed cost is specifiable as

$$Y = mQ - F, \tag{9-3}$$

where F stands for the fixed cost of production. Profits for each market area can thereby be obtained, as in equations (9-4i) below:[(II)]

$$Y_t = 6K_{*t}^2 \left(\frac{(\sqrt{3})b}{2} - \frac{(\sqrt{3})m_t}{2} - 0.7969K_{*t} \right) m_t - F; \tag{9-4t}$$

$$Y_s = 8K_{*s}^2 \left(\frac{b}{2} - \frac{m_s}{2} - 0.3848K_{*s} \right) m_s - F; \tag{9-4s}$$

$$Y_h = 12K_{*h}^2 \left(\frac{b}{2\sqrt{3}} - \frac{m_h}{2\sqrt{3}} - 0.2027K_{*h} \right) m_h - F; \tag{9-4h}$$

3. For example, in the case of a square market $\rho = 4$, $\pi/4$ is the upper limit for θ, and $K_0 = K_*/\cos(\pi/4)$ is the upper limit for K.

$$Y_c = 2\pi K_{*c}{}^2 \left(\frac{b}{2} - \frac{m_c}{2} - \frac{K_{*c}}{3}\right) m_c - F; \tag{9-4c}$$

where subscripts t, s, h and c respectively represent triangular, square, hexagonal, and circular market areas.

The profit-maximizing prices for the hexagon *(h)* and circle *(c)*, and also for the triangle and square, are easily obtained by taking partials of (9-4) with respect to m. We record the prices of the hexagon and circle as (9-5h) and (9-5c):

$$m_h = \frac{b}{2} - 0.3509K_*; \tag{9-5h}$$

$$m_c = \frac{b}{2} - 0.3333K_*. \tag{9-5c}$$

Significantly, the coefficients of K_* denote the freight-absorption rates for the given market area shapes. Substituting the m values back into (9-4) *and* taking the derivatives with respect to K_* yield the values of K_* which maximize profits—call them $K_*{}^*$. The $K_*{}^*$ for the hexagon and circle are given in Table 9-1. This table shows that as the market area approaches a circle, maximum profits increase. Moreover, if costs and demand are related as indicated in (A),

$$0.1098b^4 < F < 0.1104b^4, \tag{A}$$

profits would be positive in the circular market area, but negative for hexagonal market areas. According to Mills and Lav, this possibility disproves the theorem that market areas must be space-filling in order to

Table 9-1 *Mills and Lav on alternative market areas and profits under monopoly*[a]

	HEXAGON	CIRCLE
Profit-maximizing market size $K_*{}^*$	$0.7121b$	$0.7501b$
Maximum profit Y^*	$0.1098b^4 - F$	$0.1104b^4 - F$

[a] In the Mills and Lav original table, profit-maximizing $K_*{}^*$ for the hexagon, for example, is recorded as $-0.7121(c_3/c_4)$, where $c_3 = -t(b - ac)$ and $c_4 = at^2$. Readers who have studied their paper should note that, pursuant to the use of symbols in this book and other writings, we have reversed their b and a parameters and are using c in place of their marginal cost symbol k. Our values for K_* can then be derived under the assumptions $a = 1$, $c = 0$, and $t = 1$. The system we are describing above is a "simple" transform of that of Mills and Lav.

yield industry equilibrium. It therefore follows that hexagons (and any other regular polygon with less than six sides) cannot yield positive profits when costs are at the level defined in (A), even though competition has caused the circle to fall below its profit-maximizing size $0.7501b$.

Mills and Lav recognize that (A) is a trivial exception to Lösch's theorem, since it requires costs and demand to be such that hexagonal market areas never result in positive profits. More generally than (A), Mills and Lav also claim that there are identical hexagon and circle sizes (i.e., identical K_*'s) smaller than the profit-maximizing hexagon size, in each of which profits would be positive, *but where profits in the inscribed circle would still be greater.* They do this by substituting (9-5h) and (9-5c) into (9-3) to obtain

$$Y_h(K_*) = 0.4269K_*^4 + 1.2160bK_*^3 + 0.8600b^2K_*^2 - F; \qquad (9\text{-}6h)$$

$$Y_c(K_*) = 0.3491K_*^4 + 1.0472bK_*^3 + 0.7854b^2K_*^2 - F. \qquad (9\text{-}6c)$$

Then setting $Y_h = Y_c$, three values of K_* are derived which satisfy the equality of profits criterion, 0, $0.7075b$, and $1.4624b$. Between 0 and $0.7075b$, $Y_h > Y_c$, between $0.7075b$ and $1.4624b$, $Y_h < Y_c$, and for values of $K_* > 1.4624b$, $Y_h > Y_c$. Most important, for values of K_* between $0.7075b$ and $0.7121b$ (call them K_{**}) $Y_h < Y_c$. It is clear that competitive impacts could press market areas to sizes K_{**} which are less than $0.7121b$ but greater than $0.7075b$. Thus Mills and Lav assert that profits in circles could be greater than profits in hexagons even under conditions where competition would permit positive profits in hexagons or their inscribed circles of sizes somewhat less than the most profitable hexagon size $0.7121b$. *However, a fundamental economic error holds for the mathematics set forth by Mills and Lav.*

Greenhut and Ohta

The profit-maximizing market size K_*^* for the triangle, i.e., the *shortest* distance from the seller to the perimeter, is $0.5434b$, according to Mills and Lav (see Table 9-2). But this implies that the (profit-maximizing) *longest* distance from the seller to the perimeter, call it K_0^* as compared with K_*^*, is $2 \times 0.5434b$. This value is clearly greater than the profit-maximizing K_*^* for the circle.[4] Because the same profit-maximizing price $0.2500b$ applies to the triangle as to the circle (see [4], Appendix to Chapter 6 above, and Table 9-2), it follows that $m^*(= 0.25b) + K_0^*(= 1.0868b)$ inevitably involves negative demand at the most distant

4. By rules of elementary geometry, we know that the shortest distance from the center to any side of a regular triangle is equal to half the longest distance from the center to any boundary point.

Table 9-2 *Mills and Lav on alternative equilibrium solutions under spatial monopoly*

	TRIANGLE	SQUARE	HEXAGON	CIRCLE
Profit-maximizing market size K_*^*	$0.5434b$	$0.6534b$	$0.7121b$	$0.7500b$
Maximum profit Y^*	$0.0959b^4 - F$	$0.1067b^4 - F$	$0.1098b^4 - F$	$0.1104b^4 - F$
Profit-maximizing mill price m^*	$0.2500b$	$0.2500b$	$0.2500b$	$0.2500b$

point(s) of the triangular market. The Mills-Lav finding, that $K_*^* = 0.5434b$ in the case of a triangular market area, violates the profit-maximization rules which support it, since the buyers farthest from the seller in the triangle must purchase negative quantities (see Fig. 9-1).

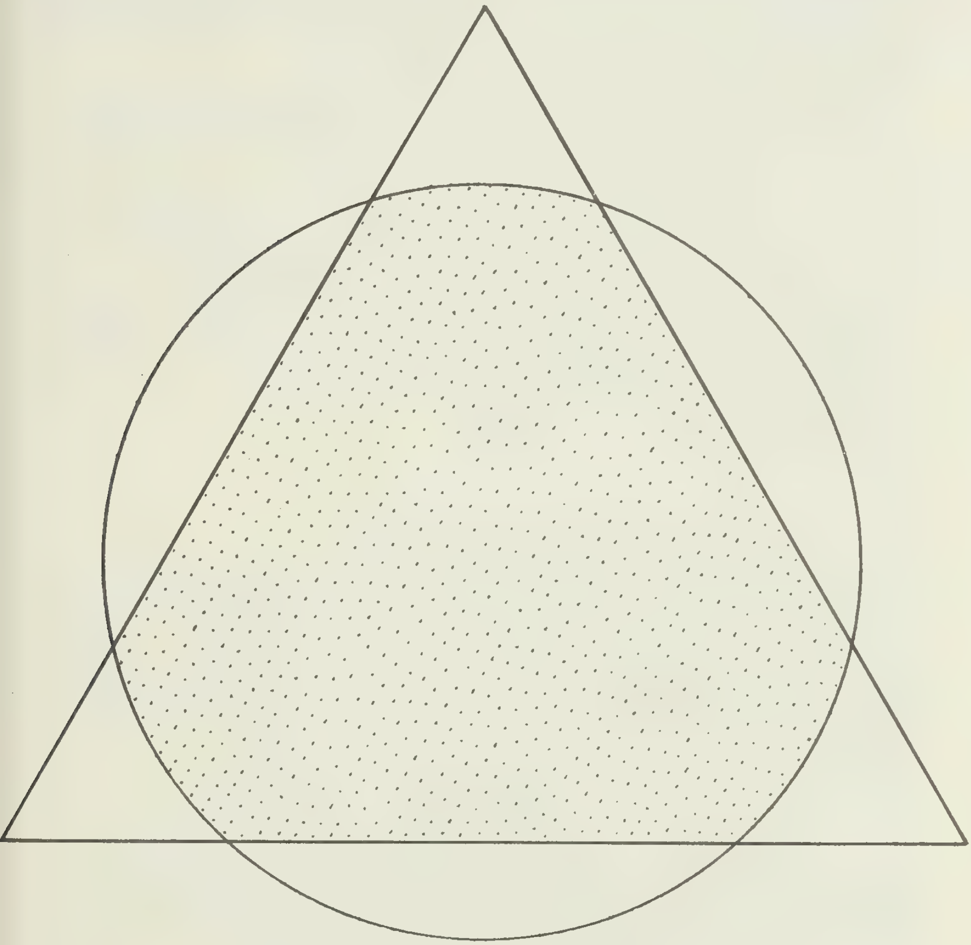


Fig. 9-1 Eroded triangular market area.

In other words, the profit-maximizing triangle of Mills and Lav exists only in the mind's eye [2]. (Incidentally, it will be shown later that the true $K_0{}^*$ in the triangle *is* $0.7500b$, while the true $K_*{}^*$ *is* $0.3750b$, not the $0.5434b$ which Mills and Lav obtained.)

Demand is negative at distant points of the Mills-Lav triangle of size $K_*{}^* = 0.5434b$ and $K_0{}^* = 1.0868b$. Manifestly, the price $m^* = 0.25b$ plus the distance $K_0 = 0.75b$ sums to the intercept value b, a value for $K_0{}^*$ which accordingly sets the distance limit of a spatial market, given $m^* = 0.25b$. Simply put, Mills and Lav overlooked the nonnegative quantity requirement in economic theory. A constraint, therefore, is needed for their system, since $K + m \leq b$ cannot be met by any $K > 0.75b$ when price is equal to $0.25b$. The Mills-Lav *actual* triangular market space is illustrated by the shaded area in Fig. 9-1; in other words, the resulting market area is not a strict triangle.[5]

The *longest* distances from the firm to the perimeter of the square and to the perimeter of the hexagonal markets are respectively $0.6534\sqrt{2}b$ and $0.7121\ (2/\sqrt{3})b$.[(iii)] These distances also are clearly longer than the profit-maximizing distance for the circle market. Hence, the square and hexagon cannot exist in the formal sense of the term *if* mill price is equal to $0.25b$. Small wonder that some large-size polygons were found to be less profitable than their much smaller-size (inscribed) circles.

In order to obtain the true profit-maximizing size of any regular polygon, the demand constraint mentioned above, i.e., $K \leq b - m$ must be imposed. Since the farthest point from the firm selling over any polygon or circle is the point where demand vanishes for any given mill price m, the following equality is required:

$$K_{0i} = b - m_i, \quad i = t, s, h, c, \tag{9-7}$$

where we use, as before, subscript i to refer to the shape of market area. Thus, for example, K_{0t} stands for the longest distance from the seller to the perimeter of a triangular market. Pursuant to a simple geometric manipulation of market size, the shortest distance K_{*i} (which in each regular polygon is in some ratio to the longest straight-line distance to a boundary point) can be expressed in terms of K_{0i} and hence in terms of $b - m_i$. Thus:

5. One might argue that $K_{*t} = 0.5434b$ is the profit-maximizing size of the *nominal* triangle market. If, however, interest lies in nominal polygonal markets, then profit-maximizing K_{*t} must be equal to $0.75b$, i.e., equal to the profit-maximizing K_{*c}, and in effect a regular triangle superscribes the profit-maximizing circle. More generally, the maximum profit for any *nominal* polygon would be exactly the same as that of the circle market. This is not the case in the Mills and Lav model because they mechanically subtracted *negative* profits from the firm's actual profit and therefore derived, for example, $K_{*t} = 0.5434b$.

$$K_{*t} = \frac{1}{2}(b - m_t); \tag{9-8t}$$

$$K_{*s} = \frac{(b - m_s)}{\sqrt{2}}; \tag{9-8s}$$

$$K_{*h} = \frac{\sqrt{3}(b - m_h)}{2}; \tag{9-8h}$$

$$K_{*c} = b - m_c. \tag{9-8c}$$

Profits in equations (9-4i) under the corresponding constraints (9-8i) can be maximized by differentiation with respect to K_{*i} or m_i. This gives

$$K_{*t}{}^* = \frac{3b}{8}, \quad m_t{}^* = \frac{b}{4}, \; Y_t{}^* = 0.0740b^4 - F; \tag{9-9t}$$

$$K_{*s}{}^* = \frac{3b}{4\sqrt{2}}, \quad m_s{}^* = \frac{b}{4}, \; Y_s{}^* = 0.0961b^4 - F; \tag{9-9s}$$

$$K_{*h}{}^* = \frac{3\sqrt{3}b}{8}, \quad m_h{}^* = \frac{b}{4}, \; Y_h{}^* = 0.1074b^4 - F; \tag{9-9h}$$

$$K_{*c}{}^* = \frac{3b}{4}, \quad m_c{}^* = \frac{b}{4}, \; Y_c{}^* = 0.1104b^4 - F. \tag{9-9c}$$

The profit-maximizing values of K_{*i} and m_i resulting from the nonnegativity requirement may be contrasted with those recorded in Table 9-2 (see Table 9-3). The profit-maximizing K_{*i} and Y_i (like those which Beckmann might have obtained if he pursued his own system of relationships to this end [1, pp. 45 ff.]) are all smaller because of the nonnegativity constraint than the values obtained by Mills and Lav *without* the constraint. This follows because they actually formulated market areas which only approximated the hexagon, the square, and the triangle rather than a strict polygon. In other words, our results reflect (and provide) regular polygons, each being inscribed in the profit-maximizing circle. Interestingly, the profit-maximizing prices are all the same, being equal to $b/4$ even under the nonnegativity constraint.

Most important for the question of profit-maximizing sizes is the following. If the profit-maximizing K_* for the hexagon market, i.e., $3\sqrt{3}b/8$, is substituted into the profit function for the circle market (expressed in

Table 9-3 *Alternative equilibrium solutions under spatial monopoly: our solutions compared with Mills and Lav's*[a]

	TRIANGLE	SQUARE	HEXAGON	CIRCLE
Profit-maximizing market size K_*	$\frac{3b}{8}$ (or $0.3750b$), not $0.5434b$	$\frac{3b}{4\sqrt{2}}$ (or $0.5303b$), not $0.6534b$	$\frac{3\sqrt{3}b}{8}$, (or $0.6495b$), not $0.7121b$	$\frac{3b}{4}$ $\frac{3b}{4}$
Maximum profit Y^*	$0.0740b^4 - F$, not $0.0959b^4 - F$	$0.0961b^4 - F$, not $0.1067b^4 - F$	$0.1074b^4 - F$, not $0.1098b^4 - F$	$0.1104b^4 - F$, $0.1104b^4 - F$
Profit-maximizing mill price m^*	$\frac{b}{4}$ $\frac{b}{4}$	$\frac{b}{4}$ $\frac{b}{4}$	$\frac{b}{4}$ $\frac{b}{4}$	$\frac{b}{4}$ $\frac{b}{4}$

[a] The upper values recorded in each tier are our values. The lower values in each tier are those of Mills and Lav.

terms of the same K_*, the profit for the circle market turns out to be less than that for the hexagon, *contrary to the Mills-Lav assertion.*[6] Thus, there is no K_{**}, because when $K_{*c} = K_{*h} = 3\sqrt{3}b/8$, $Y_h = 0.1074b - F$ and $Y_c = 0.1006b - F$; hence $Y_h > Y_c$. More generally, this implies that the profit curve for the hexagon intersects that of the inscribed circle (from above) at a point where Y_h is *decreasing* and not increasing.[7] This relation alone disproves their fundamental assertion and (re)establishes Lösch's theory of spatial equilibrium. But more important than this, it indicates that the general relationships between prices and market area sizes and shapes under spatial competition must yet be determined. Moreover, a theory of the evolution of market areas during the period when the system moves towards and reaches the spatial zero-profit equilibrium is needed. These are the main tasks assigned to Chapter 10.

IV. *Preview of Chapters 10 and 11*

The last two chapters of this book are concerned essentially with several related questions; in particular the market shapes and sizes of

6. It is worth noting that in the absence of the nonnegativity condition the general shape of the profit curve in terms of K_{*i} is a W-shaped curve which is unbounded from above. Under the nonnegativity requirement, it is intrinsically a reverse W-shaped curve. Hence, while our model guarantees an absolute maximum, the unconstrained market area model of Mills and Lav only points out a relative maximum. Let us say that since they do not impose any constraint on the domain of the function, there is no reason why their relative maximum should have become the (absolute) maximum.

7. In contrast, see [7, p. 281].

spatial monopolists and spatial competitors pricing either f.o.b. mill or discriminatorily must be determined. In the process, the thesis will be sharply established that (1) the circle cannot prevail under the specifications of a competitive equilibrium because this market area shape would, in fact, be less profitable than its alternative polygon. Then (2) it will be shown that the hexagon *is* the market area shape which, in general, yields a competitive equilibrium, including the zero-profit competitive equilibrium in economic space. At the same time (3) the correct relations between prices, profits, and spatial market areas under conditions of competition are set forth. Finally, (4) the impact of spatial price discrimination on sizes and shapes of market areas under competition, *including* the zero-profit competitive equilibrium in economic space is determined. Though these chapters refer often to Lösch's theory and the Mills-Lav counterclaim towards the end of indicating why the former theory is the correct one, the objective is not simply to reestablish one theory at the expense of another. As should be clear from previous statements, these chapters will extend market area-spatial price theory well beyond the relations scored by Lösch. They will do this by pointing out what happens to prices and profits and market area shapes under a discriminatory as well as the (nondiscriminatory f.o.b.) price system when spatial competition replaces spatial monopoly. Most important, the *spatial* equilibrium properties of firms in zero-profit competition will be (re)established in the belief that the classical welfare properties of the zero-profit equilibrium of the firm in economic space also apply (or if you will, can also be (re)established [6]). The theory of the firm will thus acquire not only a time dimension but a spatial (trading area) dimension. Manifestly, the effectiveness of any economic society has to be evaluated on both levels.

References

[1] M. Beckmann. *Location Theory.* New York, 1968.

[2] M. L. Greenhut and H. Ohta. "Spatial Configurations and Competitive Equilibrium." *Weltwirtschaftliches Archiv,* March 1973, pp. 87–104.

[3] M. L. Greenhut. *Plant Location in Theory and in Practice.* Chapel Hill, N.C., 1956.

[4] M. L. Greenhut. *A Theory of the Firm in Economic Space,* 2d printing, Austin, 1974.

[5] E. M. Hoover. *The Location of Economic Activity.* New York, 1948.

[6] A. Lösch. *Die räumliche Ordnung der Wirtschaft.* Jena, 1944.

[7] E. S. Mills and M. R. Lav. "A Model of Market Areas with Free Entry." *Journal of Political Economy,* 1964, 72:278–88.

10. Equilibrium shapes and sizes under nondiscriminatory pricing in economic space

I. *Introduction: a spatial monopolist subject to a discrete buyer distribution*

This chapter assumes f.o.b. mill pricing. (1) It argues that the circular market area cannot exist under the requirements of spatial competition, because this market area shape would in fact be less profitable than a superscribed polygon. (2) It contends that in general the hexagon *is* the market area shape which yields a competitive equilibrium in economic space. Also (3) it determines the precise relationships between prices, profits, and market areas under conditions of spatial competition.

A slightly different model is first set forth from that previously used in order to provide a general picture of monopoly price, quantity produced, profits, and other economic relationships in different-shaped market areas. Unlike the model in Chapter 9, the model used herein is related to discrete as well as continuous distributions of buyers over economic space, and the present introductory section is devoted to such formulation. Our basic model can also be generalized to include determination of the impacts of increasing spatial competition on distances and prices over the firm's market area. In fact, the inquiry in later sections of this chapter goes beyond previous investigations to determine the optimizing market area forms under conditions of varying degrees of nontrivial competition in economic space besides that of simple monopoly. But consider first a spatial monopolist selling to buyers discretely (not continuously) distributed over an economic space, as in [7]. For the sake of simplicity, assume that actually the buyers are evenly scattered over the plain in the following pattern: (a) they are so distributed that each buyer is located (as if) at the intersection point of the lines on a grid or, say, the lines on graph paper. The subject seller in turn is located in the middle of the grid. Transport routes to any buyer are assumed to follow the lines of the grid, with no short cuts (over the empty plain) available to the shipper. Though our spatial distribution is an extreme one, it

nevertheless represents the real world situation much better than the usual assumption in space microeconomics of a strictly continuous distribution of buyers over the plain. At least, intraurban transportation in many cities follows the outlines of right-angled, intersecting streets. Under present assumptions, the seller's market may be considered a square-shaped market area. As will later be apparent, conclusions derived from the present case support, with a few modifications, the analysis based upon a continuous distribution of buyers.

Other beginning conditions should be specified: For simplicity, assume that (b) marginal costs of production are zero, (c) transportation cost per unit per mile is unity, and (d) all buyers are possessed of the same linear demand curves, each of unitary (negative) slope. In accordance with these assumptions, equation system (I) is applicable. Thus:

$$Q = 4\sum_{K=0}^{K_0} K(b - m - K); \qquad (10\text{-I-1})$$

$$Y = mQ - F; \qquad (10\text{-I-2})$$

$$K_0 = b - m; \qquad (10\text{-I-3})$$

$$\frac{dY}{dK_0} = 0 \quad \text{or} \qquad (10\text{-I-4})$$

$$\frac{dY}{dm} = 0. \qquad (10\text{-I-4})'$$

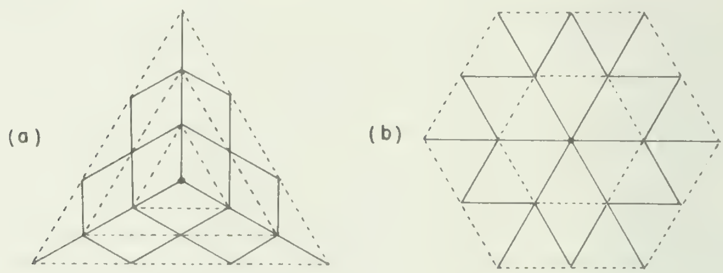
where Q stands for the output produced (and sold) by a firm, m is the mill price, b is a positive constant which establishes the price intercept of the demand function, Y is profit, F stands for fixed cost, and K represents discrete units of distance ($K = 0, 1, 2, \ldots, K_0$) from the seller's site to any buyer's site along the transportation route, K_0 standing for the distance to the market boundary. It should be recalled here that in Chapter 9 K_0 was used to represent the *longest* distance to the market boundary. However, the present model of discrete distribution of buyers does not require distinction between the shortest and the longest distances; i.e., they are the same distance. Hence, either K_0 or K_* can be employed to refer to *the distance* from the seller to the market boundary. K_0 is henceforth designated to represent this distance.

Equation (10-I-1), or more exactly the right-hand side of (10-I-1), establishes the spatial demand function. It is subject to assumptions (a), (c), and (d) and is derived from each buyer's demand schedule at each point in the space. The firm's profit, pursuant to assumption (b), is designated by

Table 10-1 *Spatial monopoly equilibrium solutions reformulated*[a]

	m^*	K_0^*	Q^*	$Y^* = m^*Q^*$
Triangle	$\frac{b}{4}$	$\frac{3b}{4}$	$3Q_0$	$\frac{3bQ_0}{4} - F$
Square	$\frac{b}{4}$	$\frac{3b}{4}$	$4Q_0$	$\frac{4bQ_0}{4} - F$
Hexagon	$\frac{b}{4}$	$\frac{3b}{4}$	$6Q_0$	$\frac{6bQ_0}{4} - F$
ρ-gon	$\frac{b}{4}$	$\frac{3b}{4}$	ρQ_0	$\frac{\rho bQ_0}{4} - F$

[a] When $\rho = 3$, we are conceiving of a triangular market. Again we must constrain the transportation system (routes) as we did in the square market case discussed above. In the present case, the (3) routes from the seller to the buyers must be at 120° angles as shown in the triangular figure below. Moreover, in the special case of a triangular arrangement, buyers must be distributed spatially in the pattern described by the triangular figure below, i.e., at the intersecting points of the transport routes depicted by the unbroken lines drawn in the figure. (Observe that the broken lines depict the underlying triangular shape of the seller's market.) When $\rho = 6$, we conceive of a hexagonal market. Transport routes to a given buyer (somewhat similar to the square market case) must follow the pattern described in the hexagon-shaped figure depicted below by the dotted lines.



equation (10-I-2). Either (10-I-4) or (10-I-4)′ is the profit-maximizing condition subject to (10-I-1) through (10-I-3). In fact, equations (10-I-4) and (10-I-4)′ are equivalent maximizing conditions because via (10-I-3) $dK_0 = -dm$, and hence $dY/dK_0 = -dY/dm$. Thus, $dY/dK_0 = 0$ if and only if $dY/dm = 0$. Specification (10-I-4), however, is preferred in the present model.[1]

The assumed discrete spatial distribution provides the seller with a square-shaped market area as the market is extended unit by unit (block by block). But this area can be generalized to the general polygonal market. All that is required is the substitution of (10-I-1)′ for (10-I-1):

1. The equations (10-I-4) and (10-I-4)′ are not identical requirements – in a technical sense. Actually, we cannot take the derivative of Y with respect to K_0 since it is assumed to be a discrete variable. For this reason, (10-I-4) is the precise statement of the limit to a spatial monopolist market area, not (10-I-4)′.

$$Q = \rho \sum_{K=0}^{K_0} K(b - m - K), \tag{10-I-1)'}$$

where ρ stands for the number of sides of the polygon. The solutions for the profit-maximizing m and K_0 under the monopoly model are obviously independent of ρ. The entire solutions set pursuant to the discrete case are recorded in Table 10-1, with explanation of particular derivations attached to the table. The respective m's, K_0's, Q's, and Y's for each shape of market area are derivable as follows:

$$Y = \rho m \sum_{K=0}^{K_0} K(b - m - K) - F; \tag{10-I-2)'}$$

$$\frac{dY}{dm} = 0. \tag{10-I-4)'}$$

Via (10-I-b), the summation formulae

$$\sum_{K=1}^{K_0} K = \frac{K_0(K_0 + 1)}{2} \quad \text{and}$$

$$\sum_{K=1}^{K_0} K^2 = \frac{K_0(K_0 + 1)(2K_0 + 1)}{6},$$

and (10-I-3), we then obtain $m^* = b/4$. In turn, since we know that $m^* + K_0 = b$, $K_0^* = (\frac{3}{4})b$. The Q_0 (of $Q^* = \rho Q_0$) is derived via

$$Q_0 = \sum_{K=0}^{K_0} K(b - m - K)$$

for the discrete case, and

$$Q_0 = 2\rho \int_0^{\pi/\rho} \int_0^{K_*/\cos\theta} \{b - (m + K)\}\, K dK\, d\theta,$$

for the continuous case.

The implication of our model is clear. The table indicates that the greater the number of sides of the polygon (i.e., the more the polygon approaches a circle), the more profitable is the monopoly.

II. *Introduction: a spatial monopolist subject to a continuous distribution of buyers*

Basically the same conclusions can be derived from the alternative assumption of a continuous buyer distribution in place of a discrete dis-

tribution. The model which follows is a modified (revised) version of that of Mills and Lav [17] and Beckmann [2], just as the finite distribution model is a modified version of the earlier one by Greenhut [7]. The point of departure from the Mills-Lav model is that our framework is constrained by the nonnegativity condition (10-II-3) below, while theirs is not.[2]

The following equation system (II) is an exact counterpart of equation system (I). The differences in assumptions are reflected in equations (10-II-1) and (10-II-3) below, although, as noted before, the fundamental price and distance relations which derive from the system are the same as those for the finite cases:

$$Q = 2\rho \int_0^{\pi/\rho} \int_0^{K_*/\cos\theta} \{b - (m + K)\} K \, dK \, d\theta \quad \text{for} \quad \infty > \rho \geqslant 3 \qquad (10\text{-II-1})$$
$$= \int_0^{2\pi} \int_0^{K} \{b - (m + K)\} K \, dK \, d\theta \quad \text{for} \quad \rho = \infty;$$

$$Y = mQ - F; \qquad (10\text{-II-2})$$

$$K_0(K_*) = b - m; \qquad (10\text{-II-3})$$

$$\frac{dY}{dm} = 0. \qquad (10\text{-II-4})$$

where $K_0(K_*)$ stands for the distance to the boundary point farthest from the seller in any polygonal market, that distance being some definite function of K_*, the distance to the nearest boundary point from the seller.

The difference between (10-I-1) and (10-II-1) should be manifest. The difference between (10-I-3) and (10-II-3), however, requires explanation. Specifically, since the present model assumes a continuous distance variable K and correspondingly a set of continuous alternative routes, there is a direct relationship and advantage in distinguishing here between the minimum distance K_* to the market boundary and the maximum distance K_0 to the market boundary, given a polygonal market area shape. (Recall that there was no need to distinguish minimum from maximum distance to the market boundary in the cases of discrete distributions of buyers discussed above. We simply dealt in terms of integer units of distance up to the greatest distance.) For the case of a continuous

2. And cf. K. G. Denike and J. B. Parr, "Production in Space, Spatial Competition, and Restrictive Entry," *Journal of Regional Science*, April 1970, 10:54, where they claim that Mills and Lav erroneously overlook the fact that the market boundary is an implicit function of price, and that "a rounded hexagon" rather than a nonhexagonal polygon would appear under spatial competition. Unfortunately, their claim is somewhat vague in that they never mention the nonnegativity condition in their criticism of Mills and Lav, albeit their argument of rounded hexagons is well taken.

distribution of buyers, if demand has not vanished at the most distant buying point currently included in the seller's market area, the spatial monopolist would be better off expanding his sales radius at the given m. But if the numerical value assigned to quantities purchased would become negative should the seller attempt to expand his sales radius, he could not sell to this more distant portion of his market. This is the meaning of the (nonnegativity) specification (10-II-3), where any shape of market area may be conceived. The profit-maximizing firm in economic space is constrained by (10-II-3), with K_0 and K_* being related. Any maximizing solution without this constraint would fail to yield the optimal economic size market areas. As mentioned previously, the values shown in Table 10-1 apply also in general to the continuous distributions. (And see Chapter 9, Table 9-3, for other data.)

III. *Our basic model of spatial competition*

What are the impacts of free entry on the firm's prices and the size and shape of its market area? In answer, one may immediately expect that the size of a firm's market area would be reduced by the imposition of an upper bound (or more exactly, a least upper bound) on the firm's delivered price. Indeed, recall from Chapter 8 that this limit is imposed on a seller by a distant competitor who is assumed to be identical to the subject seller in every relevant respect except location. The natural price limit due to the nonnegativity condition represented by equation (10-II-3) must accordingly be replaced by an alternative exogenous constraint, as in (10-III-3) recorded later. This new upper limit to delivered price depends on the degree of competition, even though it is a parameter to the individual seller. The greater the extent of the competition, i.e., the greater the number of rivals, and/or the nearer the rival is located to the subject seller, the lower will be the seller's highest delivered price, and hence the lower will be the amount of profit. Under free entry, the upper limit to the delivered price must therefore be determined in such way as to push every firm's windfalls to zero. This means that another constraint will be required for equilibrium in economic space, namely, (10-III-5), recorded later; the other equations of (II) will be seen to remain unchanged.

The initial impacts of competition

Before recording the new equation system discussed in the preceding paragraph, and in turn setting forth the results obtained for the zero-profit equilibrium, it is desirable to describe some of the initial effects of (slight) competition in the space. In this key, note that in Table 9-3 the

profit-maximizing K_{*i} and Y_{*i} were found to be smaller in polygonal markets subject to the nonnegativity constraint than were the corresponding values obtained by Mills and Lav *without* this constraint. These smaller values are due to the fact that our results reflect (and provide for) regular polygons, each being *inscribed* in the profit-maximizing circle rather than lying partly within and without that circle. Interestingly enough, the profit-maximizing prices are all the same, equal to $b/4$, even under the nonnegativity constraint.

We further found that substitution of the profit-maximizing hexagon K_*, i.e., $3\sqrt{3}b/8 (= 0.6495b)$, into the profit function for the circle market would, *contrary to the Mills-Lav thesis,* provide smaller profits for the circle of size $K_* = 3\sqrt{3}b/8$ than it does for the profit-maximizing hexagon of this same size K_*. There are accordingly no K_*^* sizes given the nonnegativity constraint where $Y_h < Y_c$, since when $K_{*c} = K_{*h} = 3\sqrt{3}\,b/8$, $Y_h = 0.1074b^4 - F > Y_c = 0.1006b^4 - F$. Hence, for competitive size areas less than (or equal to, *or even slightly greater than*) the profit-maximizing hexagon size, Y_h is always $> Y_c$. An inscribed circle, in other words, is always less profitable than any properly constrained (economic-size) hexagon. (Compare the profit curves in Fig. 10-1(a) with the Mills-Lav profit curves of Fig. 10-1(b).)

The constraint $K + m \leqslant b$, or more precisely, $K_0 + m \leqslant b$, thus signifies that more profitable irregular polygons may always be superscribed over "competitively" reduced circles (i.e., circles whose sizes are $<0.75b$). But since a network of regular polygons (including polygons approaching a circle) must be more efficient than a network of irregular polygons, Table 9-3 indicates that any K_* falling between $0.6495b$ and $0.75b$ will involve market shapes characterized by regular polygons. These larger-than-hexagonal sizes have sides which decrease in number towards 6 as $K_* \to 0.6495b$. It is only to the extent that this "happening" occurs in economic space that competitive market areas would be scored by polygons other than the hexagon.[3]

A tendency to minimize the "unserviced" space thus prevails in the competitive free enterprise system. Most fundamentally, there is no competitively reduced market area below the profit-maximizing hexagonal size for which the inscribed circle is more profitable than the superscribed hexagon. Only the trivial case (where cost is such as to require monopolistic or quasi-monopolistic market areas) rules out the hexagon. The economic constraint $K_0 + m \leqslant b$ signifies that the profit curve for the hexagon intersects that of its inscribed circle (from above), viz., at a point where Y_h is *decreasing,* not increasing. Put differently, $dY_h/dK_{*h} <$

3. If $0.1104b^4 > F > 0.1074b^4$, circles would obtain under monopoly and eroded circles would obtain under competition, but never would circles hold in the presence of competition, even in the trivial case.

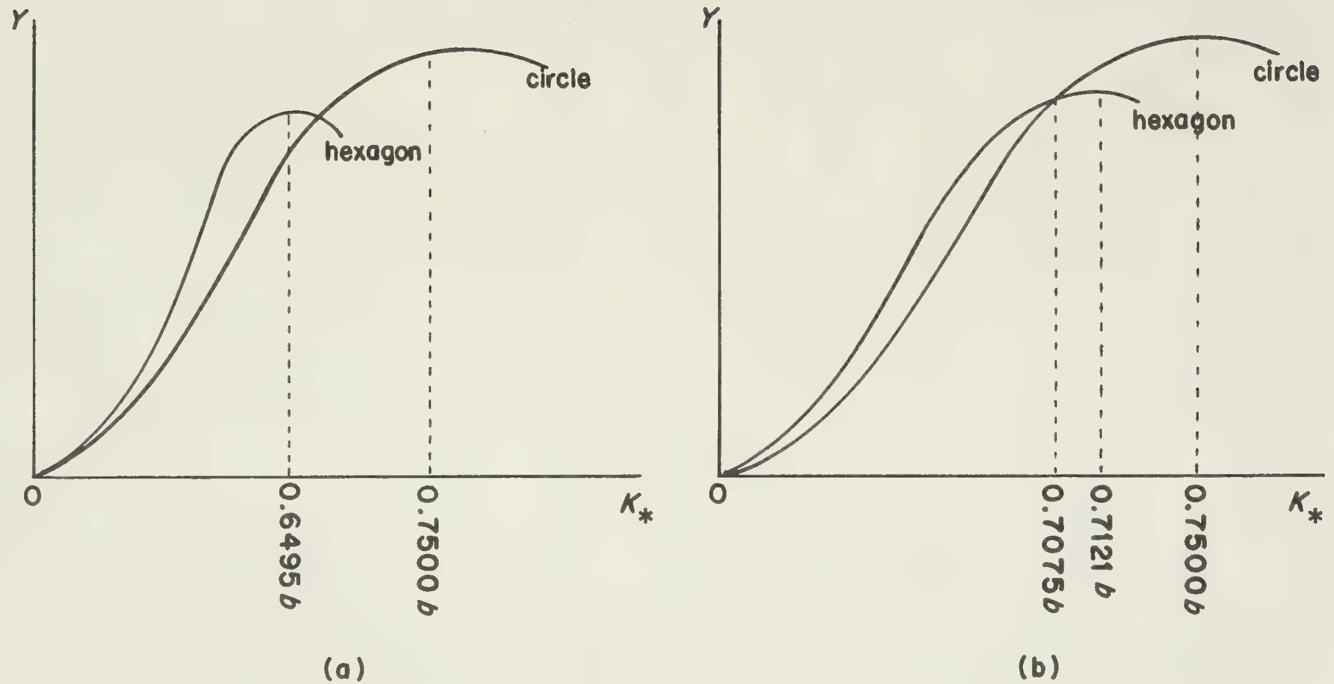


Fig. 10-1 Profit curves under alternative market area shapes.

0 when K_{*h} is of such greater size than $0.6495b$ as to yield $Y_h = Y_c$.

Under conditions of free entry, and costs $F \leqslant 0.1074b^4$, the hexagon therefore arises. But does the hexagon continue to prevail all the way to the limit of any $0 < K_* < 0.6495b$? This query indicates that there remains the rather weighty question, "What are the relationships between prices and market areas under (increasing) spatial competition?"[4] More generally, there remains the requirement of setting forth the relationship between prices and market areas as the free enterprise system moves toward (and reaches) the spatial zero-profit equilibrium, a subject too long untouched in the literature. Equation system (III) below is designed for this purpose.

Impacts of substantial competition

Equation system (III) may be referred to as the system of equations applicable to a competitive spatial equilibrium under any particular market shape. The equations are as follows:

$$Q' = 2\rho \int_0^{\pi/\rho} \int_0^{K'/\cos\theta} \{b - (m' + K)\} K \, dK \, d\theta; \tag{10-III-1}$$

$$Y' = m'Q' - F; \tag{10-III-2}$$

$$K'_0(K'_*) = p' - m', \quad b > p' > m'; \tag{10-III-3}$$

$$\frac{\partial Y'}{\partial m'} = 0; \tag{10-III-4}$$

$$Y' = 0; \tag{10-III-5}$$

where p' is the exogenously given maximum delivered price for our seller, and all endogenous variables are now primed, since they are constrained by competition to depend on p'.[5] Thus, equations (10-III-1) through (10-III-4) are exact counterparts of equations (10-II-1) through

4. We are using the words "weighty question" in deference to another study the present authors are undertaking, which is designed to show that market areas of sizes $K_* \geqslant 0.6495b$ are empirically unlikely in our time under conditions of spatial competition. Hence, in the short run as well as the long run, we may expect hexagonal market areas to prevail in a spatially competitive free enterprise economy.

5. It must be noted that p' is treated as an unknown variable in our new system of equations, albeit competitive sellers consider it to be a parameter exogenously given to them. Therefore, the differentiation in (10-III-4) should now be interpreted as partial, not total, differentiation, as was the case for (10-II-4). Note also that when p' is at its upper limit b, the system converges to that of spatial monopoly. When p' is at its lower limit m, the system converges to the spaceless one, as is readily evidenced by the condition that when the delivered price is m, $K_0 = 0$.

Table 10-2 *Alternative spatial equilibrium values under free entry*[a]

	m'^*	K'^*	$K'_0{}^*$	p'	Q'^*	Y'^*
Triangle	$0.2227b$	$0.2913b$	$0.5826b$	$0.8053b$	$\frac{0.05}{0.2227}b^3$	0.0000
Square	$0.2002b$	$0.3409b$	$0.4820b$	$0.6822b$	$\frac{0.05}{0.2002}b^3$	0.0000
Hexagon	$0.1873b$	$0.3744b$	$0.4323b$	$0.6196b$	$\frac{0.05}{0.1873}b^3$	0.0000
Circle	$0.1789b$	$0.4008b$	$0.4008b$	$0.5797b$	$\frac{0.05}{0.1789}b^3$	0.0000

[a] The values in the table were derived under the assumption that $F = 0.05b^4$. The mill price m'^* recorded in the table is obtained first by determining p' and K'_* via equation (10-III-a) listed below, and applying other equations, explained next.

Take

$$p'_i = m'_i + K'_{0i}(K'_{*i}), \quad i = t, s, h, c. \qquad \text{(10-III-a)}$$

We know that the maximum distances for the various polygons are each in a fixed relation to the radius of the inscribed circle. Therefore m'_i can be expressed in terms of K'_{*i}, given p'_i. From (10-III-1) and (10-III-2), the profit function can then be specified in terms of m_i and K'_{*i}. This gives us $Y'_i = Y_i(m'_i, K'_{*i}; p'_i, F)$. Then applying the profit-maximizing condition (10-III-4) subject to (10-III-a) and (10-III-5) for each shape of market area yields three equations in three unknowns, namely, Y'_i, K'_{*i}, and p'_i and in turn the data recorded in the table above. Further details are provided in Appendix I to this chapter.

(10-II-4), which represented the monopoly model. The equilibrium solutions for the system of competition in economic space are summarized (along with the explanation of results) in Table 10-2.

It is apparent in Table 10-2 that a hexagon superscribed over the circle of size $K'_{*c}{}^* = 0.4008b$ would offer profits to the firms. And such superscribing is possible because there are buyers who otherwise would not be purchasing any of the product. But profits for the hexagonal seller will promote entry. In turn, the entry of new rivals will compress the size of the hexagon while allowing the firm to increase its mill price slightly. It follows that the sales of the individual firm, now selling over a smaller hexagon than the one which would superscribe the circle of size $K'_{*c}{}^* = 0.4008b$, falls partly within and partly without that circle. The ultimate profit position of the hexagonal seller must be zero in the competitive equilibrium, as the higher mill price he is able to charge compared to the seller in circle size $0.4008b$ is offset by the smaller quantity of sales available to this firm.

To appreciate these relations fully, suppose no entry and a given technology are assumed. As was shown above, a firm would then sell in a circle market area, since this market provides the greatest profit. However, Table 10-2 reveals that free entry forces the maximum delivered price (i.e., the price charged at the periphery of the market) down

to $p'_c = 0.5797b$, as the circle shrinks to the size $K'_{*c}{}^* = 0.4008b$. An initial competitive zero-profit equilibrium would now exist. The size $0.4008b$ is relevant under the circular market shape *and* the further contradictory requirement (or expectation) that demand vanishes in all directions if the price p'_c rises above $0.5797b$. Of course, this market shape cannot exist, because *there are empty spaces which are profitable* (and to which sales may be made) *at delivered prices greater than* $0.5797b$. The firm will, therefore, begin to sell over market segments lying outside of the circle, for example, over a hexagon (or outline of a hexagon) market whose size $K'_{*h} = K'_{*c}{}^* = 0.4008b$. Indeed, given the circle's size $K'_{*c}{}^*$ for a hexagon, i.e., $K'_{*h} = 0.4008b$, the mill price $m'_h = m'_c = 0.1789b$ would be sufficient to provide positive profit.[6]

The positive profit under $m'_h = m'_c = 0.1789b$ and $K'_{*h} = 0.4008b$ promotes entry. After relocation and price adjustments, equilibrium eventuates when $K'_{*h}{}^* = 0.3744b$ and the delivered price limit $p'_h = 0.6196b$ obtains (see Table 10-2). These are the final equilibrium values under free entry. The square (or triangle) market of smaller size than the hexagon $K'_{*h}{}^*$ involves higher maximum delivered prices as well as higher mill prices than those in hexagonal markets and hence cannot prevail in the presence of competition.

It is significant to observe that $K'_{0h}{}^* = 0.4323b$ when $K'_{*h}{}^* = 0.3744b$. Table 10-2 also indicates that the mill (f.o.b.) price m'^* is $0.1873b$. At the profit-maximizing limit when competitive entry occurs at a distance, $m'^* + K'_{0h}{}^*$ equals $0.6196b$. Significantly, the m'^* and $K'_0{}^*$ are higher for the hexagon than the circle, though $K'_{*h}{}^*$ is smaller. The hexagon thus falls partly within and partly outside of the circle. It is finally manifest that the number of firms per square mile is greater under the hexagon, since $K'_{*h}{}^* < K'_{*c}{}^*$ and total output is greater;[7] the equilibrium sales of the individual firms must, however, be less, since $Y'_h{}^* = Y'_c{}^* = 0$ and $m'_h{}^* > m'_c{}^*$.

Impacts of spatial competition in general

The relations applicable to the hexagon and the circle help establish the basic picture of what transpires under spatial competition. This

6. The exact profit value is $0.0037b$. At this point, i.e., at $K_{*h} = 0.4008b$, according to our computer results, $\partial Y_h/\partial K_{*h} < 0$ prevails. Hence, given p'_h, the firm must be still better off if it reduces its market size K_{*h}. Note, further, that the market area shrinkage under the maximum delivered price p'_h implies necessarily a rise in mill price. And the rise in mill price is in turn attributable to the extra demand which exists for hexagonal market areas compared to the circle. In contrast, a reduction in p'_h due to free entry implies a reduction in mill price m'_h. This effect, however, would be negligible because a reduction in p'_h also implies shrinkage of the size K_{*h}, which in turn implies a rise in mill price. The net effect, therefore, is that the final equilibrium mill price $m'_h{}^*$ is slightly higher than the initial or temporal equilibrium mill price m'_c, as is evident in Table 10-2.

7. In fact, it is greater even at $K_{*h} = K_{*c}$, i.e., even before new entry, because empty spaces can be filled under the hexagon.

picture reveals that hexagonal market areas provide greater demands. Hence, delivered price $p'_h{}^*$ to the most distant market points of the hexagon may be (and is) greater than that of the circle, *ceteris paribus.* Correspondingly, the mill price m'^* of the firm may be (and is) greater when it sells over the hexagonal market area compared with the circle, while its sales total is less because of the greater number of sellers able to enter and remain in the industry, *ceteris paribus.* The classical relation of mill price being less the greater the maximum distance over which a firm sells (see Greenhut [10, app. to chap. 6]) does not apply since that situation presumes spatial monopoly pricing (i.e., the derivation of the profit-maximizing mill price under no constraints on profits from competition). The present situation involves a zero-profit constraint on the size of the hexagonal market, and this constraint is related in turn to a price constraint which is more restrictive of the circular than of the hexagonal market form. It is more restrictive because it requires both the maximum distant point of the market area and the basic mill price to be reduced more sharply in the case of the circle than their counterpart values in the hexagon (see Table 10-2). Given any nontrivial levels of competition in space, the hexagon becomes the equilibrium market area shape.

It deserves repeated emphasis that the forces promoting localization or dispersion of industries and their effects have been generally ignored. But see [10] and [7, chap. 6]. Emphasis herein has been on the market area relationships prevailing among spatially differentiated firms. Nevertheless, it can readily be shown [10, chaps. 12 and 13] that spatial competition (from within and from outside the market area of a firm) proves to be an effective force leading to productive efficiency in the long run. The final result $Y_i = 0$ (in the sense of Y standing for economic surplus) must apply, a condition which holds regardless of the number of firms that may happen to be localized at the center of any given market area.

Summary

The main market area conclusions drawn to this point of our analysis may now be summarized with advantage. In this context, recall that Lösch contended that the hexagon is most like the circle, and hence it would minimize distances and maximize demand in comparison with polygons of fewer sides. At the same time, the hexagon fills up empty spaces which the circle and polygons of more sides than the hexagon cannot do. Compared with the circle, our preceding analysis has also proved, the hexagon would be more profitable than its inscribed circle in the competitive equilibrium. The Mills-Lav thesis that there are in-

scribed circles of less than profit-maximizing size which could be more profitable than hexagons violates requisite economic constraints.

Among more specific findings, it has been shown that competition in economic space elicits a unique correspondence between prices and distances which provides the hexagon with higher prices and greater profits and distances than would the counterpart relevant circle. Moreover, the hexagon provides sufficiently greater sales and profits and lower mill prices than the triangle or square. The number of independent existences and the number of consumer satisfactions are both maximized by this market area form.

It is thus the case that what the triangle of given area size can do as a market area shape, the square market shape of the same area size can do better. What the square of that particular size can do, the circle of that same size can do better. But what the circle of that size can do, the hexagon (of greater area size) does even better. A network of hexagons blankets the space, increases profits, and, as indicated above, maximizes the number of firms existing over the space. It follows that sales and consumer satisfactions are maximized by a system of hexagons. Elsewhere, it is shown [10, chap. 8] (and see Appendix II to the present Chapter 10) that these spatial properties have a corresponding feature which involves production at least cost. But productive efficiency is, of course, another matter, whose analytical framework extends well beyond that possible in a book devoted to determining the influence of prices and profits on the market area of firms subject to competitive (free) entry of rivals at any place in economic space.

IV. *Some suggestive empirical data*

It may be desirable at this point to go outside of the field of "pure" theory to see if some applications are possible which might shed light on the actual market area sizes prevalent in the (competitive) economy of the United States. The Mills-Lav findings will again serve as our point

Table 10-3 *Mills and Lav on alternative market areas and profits under monopoly*[a]

	HEXAGON	CIRCLE
Profit-maximizing U (i.e., our K_*^*)	$-0.7121\ (c_3/c_4)$	$-0.7501\ (c_3/c_4)$
Maximum profit	$0.1098\ (c_2^2/c_4)D - A$	$0.1104\ (c_2^2/c_4)D - A$

[a] $c_4 = bt^2$, $c_3 = -t(a - bk)$, $c_2 = (a - bk)^2/b$, $c_4 > 0$, $c_3 < 0$, $c_2 > 0$, and $c_2c_4 = c_3^2$, where t stands for the constant freight rate per mile, k for constant marginal cost, and a and b are the coefficients of the demand function, which in turn are assumed to be constant. This table uses the Mills-Lav symbols exactly as set forth by these writers and to this extent differs from Table 9-1 above.

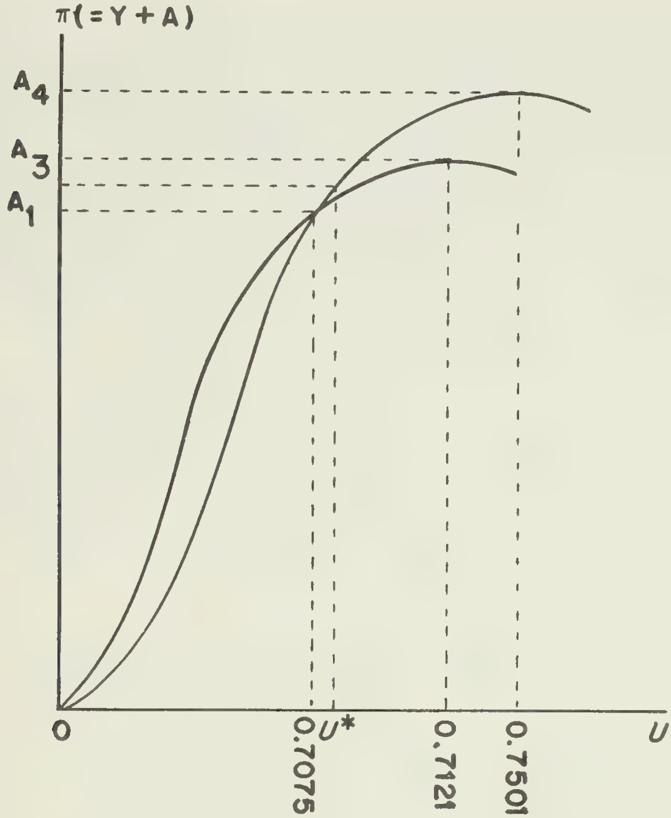


Fig. 10-2 Market area sizes.

of departure, and hence Table 10-3 and Fig. 10-2 repeat some of their findings which are relevant to our present analysis.

Revenues and costs

Mills and Lav asserted that to be viable, the hexagon must always be of size less than − $0.7075(c_3/c_4)$. Over larger relevant sizes, the inscribed circle is more profitable than the related hexagon. Fig. 10-2 embellishes on the Fig. 10-1b which depicted their findings. This new figure includes four alternative fixed-cost sums OA_1, OA_2, OA_3, and OA_4. Two "revenues minus constant marginal cost" curves OC and OH are included, respectively applicable to circular and hexagonal market areas; and equal returns are displayed for sizes zero and − $0.7075(c_3/c_4)$. One larger point of equal returns, proposed originally by Mills and Lav for the size − 1.4624 (c_3/c_4), is not included in the diagram.

Distances, freight rates, and prices

Cost and demand conditions may clearly be such that zero profits would occur coterminously with a large circular market area (e.g., with fixed cost at OA_4, Fig. 10-2). The resulting equilibrium market shape then would not be the hexagon. Alternatively, if fixed cost is OA_3 (Fig. 10-2), the hexagon is less profitable than the inscribed circle in the Mills-Lav view. Firms would, however, enter, *ceteris paribus*, compressing market areas below the spatial monopoly sizes. The resulting market area would be a regular s-sided polygon greater than the hexagon, and hence would not be space-filling.

Suppose costs are lower than OA_1 in Fig. 10-2; that is to say, consider the case where K'_* is less than or equal to $-0.7075(c_3/c_4)$. In such situations only could the hexagon prevail, according to Mills and Lav. Spatial competition thus might yield large or small market areas, depending on cost and demand conditions. Would they typically be larger than, say, size $-0.7075(c_3/c_4)$ or smaller, much smaller . . . ? In fact, would they tend to be smaller than the true critical maximum-profit size -0.6495 (c_3/c_4)?

Two pieces of information are needed before answer is given to the basic question posed above: (1) Greenhut [7, app. I], or Chapter 2 above indicates that the largest economically feasible monopoly size of market area – in the case of buyers distributed evenly over a plain – involves a transport cost to the most distant buyer(s) three-fourths the value of the price intercept less the constant marginal cost of production. Alternatively, the profit-maximizing mill price is given as $m^* = (b+c)/2 - K_0^*/3 = (b-c)/4 + c$, where c stands for constant marginal cost. (2) In Fig. 10-3, let OA_0 represent constant marginal cost. And let A_0B equal b' (where the prime on b is used here strictly to distinguish it from OB $(= b)$, the maximum price possible. A basic relation requiring proof at this time is that the transport cost for the Mills-Lav distance U (which, to recall, is measured in terms of c_3/c_4 in Table 10-3) corresponds to the K_*^* derived previously by the present authors. To establish the correspondence, consider the following.

The profit-maximizing mill price *above marginal production cost* in the case of buyers distributed over a plain is given by statement (1) above as $b'/2 - K_0^*/3 = b'/4$, where $b' = b - c$. Alternatively phrased, the mill price in Fig. 10-3 is $b'/4 + OA_0$ and the net revenue mill price (i.e., the mill price less marginal production cost) is $b'/4$. Now b' corresponds to Mills' and Lav's $[(a/b) - k]$, as can be seen from Table 10-3, where c_3/c_4 is given as $-[(a/b) - k]\ (1/t)$. Taking for the circle $0.7501 = \frac{3}{4}$, it follows that the Mills-Lav $tU = -0.7501(c_3/c_4)t$ is equivalent to the statement $tU = K_0^* = (\frac{3}{4})b'$.

At the minimum-size market area for which Mills and Lav claim any polygon with sides greater than six would prevail, i.e., where $U = -0.7075(c_3/c_4)$, freight costs are approximately *70 percent of the difference between marginal cost and the maximum price which the firm can charge.* Suppose that empirical measurements of marginal costs are often found to be of significant magnitude, such as, for example, half the *maximum price* a buyer located next door to the seller would pay for a single unit of the good. Then *even in this case* where the accountant's marginal

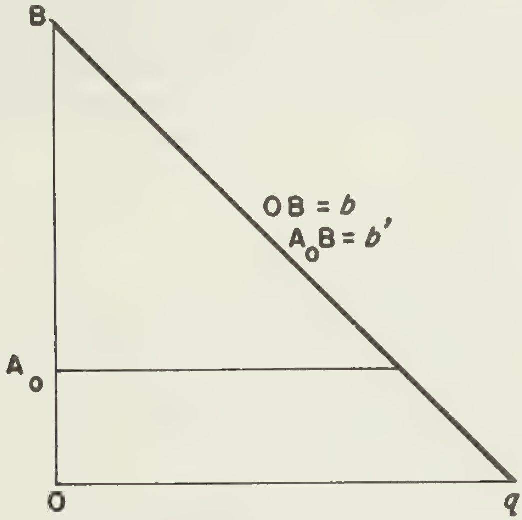


Fig. 10-3 An aid to visualizing profit-maximizing space.

production costs as compared to b are relatively high, so that only a rather small value for freight cost is possible, the freight cost to the boundary point of 70 percent of (what is in this case a comparatively small) b' would still be substantial. The applicable freight cost would amount to 35 percent of the maximum price the buyer would pay for the good. In terms of an automobile whose price intercept value (i.e., b) is \$5,600, the maximum profit trading area of a manufacturer's plant would run to the limit of \$1,960 worth of freight cost. The trading areas required for any polygon greater than the hexagon would appear to be extremely large.[8]

8. We could not find any statistics relating marginal costs (MC) to the price intercept value. However, consider the following derivation and data.

Under a negatively sloping linear demand curve and constant MC, such as was assumed by Mills and Lav, the relation $m = (b + \text{MC})/2 - K_0/3$ applies, where we are using MC in this note rather than c (as in the text above) for formula emphasis, as in (iv) below. Then:

To sum up, market areas between $-0.7075(c_3/c_4)$ and $-0.7501(c_3/c_4)$ involve freight costs which, in the present example, run from 35 to 38 percent of the *maximum price buyers would be willing to pay for a unit of the good.* This approximation (35 to 38 percent) applies to firms subject to relatively high cost of production and in turn conditions, which provide us with a rather small freight cost to consider. For hexagons of size $-0.6495(c_3/c_4)$, freight cost – in this example – would run to 32 percent of the maximum price buyers would be willing to pay for a unit of the good. Even in the present example, which provides a rather favorable cost case for nonhexagonal market shapes, market extensions of the indicated magnitude (where entry is assumed to be open) would be extremely high.[9] For a firm to have such a market area size and yet to record zero economic profits would be even more surprising.

$$m = \frac{(b + \mathrm{MC})}{2} - \frac{K_0}{3} \tag{i}$$

$$b = 2(m - \mathrm{MC}) + \mathrm{MC} + \frac{2K_0}{3} \tag{ii}$$

Substituting $K_0 = b - m$ in (ii) yields:

$$b = 4(m - \mathrm{MC}) + \mathrm{MC}. \tag{iii}$$

From $\mathrm{MR} = m(e - 1)/e$, and the profit-maximizing marginal equality, we obtain

$$\frac{\mathrm{MC}}{b} = \frac{[m(e-1)/e]}{[4m - 3m(e-1)/e]} = \frac{(e-1)}{(e+3)} \tag{iv}$$

Now in Adelman [1] two elasticities are given for typical A&P stores: $e = 1.91$ and $e = 2.33$. Using these as limits, the ratio of marginal cost to the b intercept value would be:

$$\frac{0.91}{4.91} \leq \frac{\mathrm{MC}}{b} \leq \frac{1.33}{5.33} \tag{v}$$

This yields for the MC/b ratio a set of values somewhere between 19 and 25 percent, in round numbers.

Leontief [14] provided estimates of elasticity for the U.S. Steel Company between 1927 and 1938. His estimates would yield a ratio for MC/b between 27 and 36 percent. We shall probe more deeply into MC/b ratios later in the chapter.

9. A recent study of the Bureau of Transport Economics and Statistics of the Interstate Commerce Commission (1959, p. 20) reveals that out of 54 commodity classes, the percent that "gross freight revenue" is of "wholesale value" at destination typically totaled between 1 and 5 percent. In a very *skewed* distribution, some commodity classes total between 11 and 20 percent, and two were between 21 and 40 percent. One commodity class alone (significantly furnace slag, an unlikely product type for the Mills-Lav theory as we shall later see) presented freight costs as high as 53 percent of price. But even here consider for the moment the following: (1) the ratio MC/m is > MC/b, since MC/$m = (e - 1)/e$ while MC/$b = (e - 1)/(e + 3)$; (2) transformations of Adelman's and Leontief's elasticity estimates (see note 8) point to a difference between MC/m and MC/b which averages around 30 percent; (3) this difference, in turn, suggests a similar differential with respect to freight rates and *wholesale prices* vis-à-vis the b intercept value (on the commodity classes reported in the ICC study cited above). Via the same basic price formula $m = (b + c)/2 - K_0/3$, the freight

V. *Some empirical data and theoretical expectations*

The likelihood that small market areas will prevail whenever distance is a vital cost factor may be appreciated further as a result of the following analysis. Recall in this connection that the spatial monopoly mill price is

$$m^* = \frac{b'}{2} + \mathrm{OA}_0 - \frac{K_0^*}{3}\left(= \frac{b'}{2} + c - \frac{K_0^*}{3}, \ldots \text{ where } c = \mathrm{OA}_0\right); \tag{A}$$

$$\frac{(b+c)}{2} - \frac{K_0^*}{3} + K_0^* = b. \tag{B}$$

And from (B), as observed previously,

$$K_0^* \equiv k = \frac{3b'}{4}, \tag{C}$$

where $b' = b - c$. If now $b = \$5{,}600$ and $c = 8b/10$, (A) and (C) yield the value $K_0^* = \$840$ and $m^* = \$4{,}760$. In turn, a critical k/m^* ratio stands at 18 percent. What this ratio portends warrants detailed discussion.

Table 10-4 stems from (A) and (C). The table provides relative price and distance values under spatial monopoly given assumed ratios of c to b. It can be shown that the *average* buyer (in the sense of average distance from the seller) is located at a point $(\frac{2}{3})k$, *i.e.*, $\frac{2}{3}$ the total distance

Table 10-4 *Demand, costs, and profit-maximizing distances*

	SITUATION				
RATIO	1	2	3	4	5
Costs c as a percentage of b	80	60	50	40	20
Costs c as a percentage of m^*	94	87	80	73	66
k' as a percentage of m^*	12	29	40	55	100
k as a percentage of m^*	18	43	60	82	150
k as a percentage of b	15	30	37.5	45	60

cost to the b ratio may be obtained, and as shown in Table 10-4, (4) a 53 percent freight cost to price ratio can be expected to be less than 36 percent of the value b. Again, only furnace slag out of 54 commodity classes clearly meets the percentage requirement of our example. On the basis of these and other data to be presented, it can be proposed that market areas of sellers in a competitive-free enterprise system tend to be much smaller than the profit-maximizing hexagon size.

from the seller's site to the natural boundary of his market, when his buyers are dispersed evenly (or continuously) over a plain.[1] The freight cost to this *average* buyer (call it k') is 12 percent of mill price when $c = 8b/10$. Because statistical samples of freight costs relate to sales *per buyer* (e.g., as in note 9 above), such samples therefore provide empirical values equivalent to the theoretical k' values. Recent ICC data on selected commodities (again see note 9) typically reveal extremely low k'/m^* ratios in the American economy. In fact, they were much lower than 12 percent in almost every case.

There are two basic ways or situations (under present assumptions) for a spatial monopolist selling over a profit-maximizing circle (or many-sided polygon) to experience a low k'/m^* ratio. The same two corresponding ways (or situations) would apply to a spatial competitor who is in a zero-profit equilibrium but still selling over a circular market area. *The first situation* relates to instances where the fixed costs of the firms are very high compared to the b intercept value. Unfortunately, empirical studies along this and other lines had just been initiated when this chapter was drafted, and hence only a few additional words about fixed cost and market areas can be entered later below. *The second situation* relates to the possibility that marginal costs of production are very high relative to the b intercept value (or relative to the mill price). Under this condition, a circular market area of profit-maximizing size $K_0^* \rightarrow 0.75b'$ would involve only a comparatively small freight cost. Here too empirical studies were in their very early stage of development when this chapter was drafted. Thus, only some tentative MC/b and K_0^* "likelihoods" are included below.

MC/b *ratios*

Consider initially the question how high might MC/b ratios typically be? Might they be much higher than those suggested by the elasticity estimates recorded in note 8 above? Can one assume that many very high c/m^* ratios prevail in the space economy, ratios actually as high as 94 percent? Would so high a ratio serve as an acceptable hypothesis? Could one propose that there are in fact many viable competitive firms operating with substantial c/m^* ratios and hence small mark-ups? Indeed, would a c/m^* ratio of this (94 percent) magnitude be not only quite surprising for spatial monopolists, but particularly unsatisfactory for viable competitive firms?

Table 10-4 indicates that a c/b ratio of 80 percent reflects a c/m^* ratio of 94 percent under conditions of spatial monopoly. Such a c/m^* ratio is hardly sufficient for the long-run survival of any firm subject to the risk and uncertainty of present-day business activity. (Indeed, it would surely be unlikely for a monopolist who not only is not confronted by effective

price competition but is not faced with competition in the form of market area intrusion.) If many spatial monopolies exist, one would expect ICC data to reveal substantial freight costs per buyer (i.e., substantial k'/m^* ratios). A more likely c/b ratio would appear to lie in the 50 to 60 percent region. In turn, the related spatial monopoly c/m^* ratio shown in Table 10-4 then drops to the rather high 80 to 87 percent bracket; the corresponding competitive ratio would tend to be as high as, say, 95 percent. These ratios are manifestly all too high. Moreover, the theoretical k'/m^* counterpart in such cases would run variously from 29 percent to 40 percent (see Table 10-4).

As a rule of thumb, can we not expect empirically derived c/m^* ratios to reflect some degree of spatial price competition? Actual competitive ratios (not table ratios) such as 87 or 94 percent could, accordingly, be expected to prevail quite often.[10] Adjusting these empirical ratios to reflect the monopolistic m^* intrinsic to Table 10-4 would lower the c/m^* ratio to values indicated in situations 3 to 5 in the table. Freight costs to the average buyer would then tend to run somewhere between 40 and 100 percent of m^*, again for *situations where competition exists which presumably has narrowed the trading radius of sellers to distances less than those required under conditions of spatial monopoly.* (It *is* our suggestion that an actual competitive k/m^* value of, say, 82 percent and k'/m^* ratio of 55 percent could often be approximated by a monopoly k/m^* table value of, say, 43 percent and a related k'/m^* value of 29 percent.) If empirical measures generally reflect states of competition, the expectation then follows that the k'/m^* values actually found in practice will be greater than 55 percent whenever the ravages of spatial competition still leave market areas approaching the maximum size.

Suppose empirical measures reflect states of monopoly rather than competition. It would follow that we should not expect often to obtain c/m^* values of 87 percent and k'/m^* ratios of approximately 29 percent. Very low k'/m^* ratios could be expected only under conditions where goods are sharply differentiated and freight costs are so negligible that firms literally transport their goods without distance limits.

To summarize the argument, it is proposed that the underlying c/b ratio will not be particularly high, i.e., not be greater than 80 percent, even though empirical c/m^* ratios (because of competition) are quite

10. Elementary operations with profit margins recorded by Collins and Preston [4] enable us to derive empirical c/m^* ratios; these ratios average approximately 80 percent. In only selected cases – for example, cigarettes – do the ratios run very high, up to 97 percent. And in practically each case (if not in all of these cases), the product type was one of insignificant freight cost whose market areas cover the nation or overlap. In such situations, distance is an insignificant, if not irrelevant factor. We do not include these cost data here since they would consume too much space, given our objectives, and since we believe the arguments presented above (along with the reader's intuitive expectation of what the long-run c/m^* ratios would be in a free enterprise economy) sufficiently establish our thesis. The authors will provide interested readers with the c/m^* data mentioned here.

high. Moreover, it is proposed that a limiting theoretical 12 percent value of k'/m^* would be extremely low under conditions of feasible locations at alternative points in economic space, while the likely "theoretically limiting" ratio applicable before competitive entry occurs would be much higher.

Notwithstanding arguments to the contrary, assume the prevalence of at least many very substantial c/b ratios (i.e., ratios ≥ 80 percent). In this case, the theoretical k' value would be quite low. And the ICC data recorded in note 9 above indicate that very low k'/m^* ratios actually do prevail. Could it not be that the low k'/m^* ratios do not relate to small market areas, as had been suggested up to this point? Could c/b ratios actually be high, and slight competition at a distance characterize an effective part of the American economy, with circular or many-sided polygonal market areas actually marking some spatial zero-profit equilibria? Our answer to this last set of queries is rather elementary. If c/b ratios are significantly high, industrial location will generally be attributable to the cost (not demand) factors of location [7, chaps. 4–7]. In fact, the higher the c/b ratio, the more likely it is that a nonubiquitous cost factor controlled the site-selection pattern. Depending upon the magnitude of this ratio, many firms would actually tend to locate at the same production center, or else only a few, two, or one would locate at a limited number of centers, if entry was open. Locations at other production centers depend, accordingly, on the distribution of Mother Nature's gifts. Market areas will then overlap if the very high c/b ratios are coterminous with not only low freight costs but also very low freight rates per unit of distance. Either this will be the case or else market areas will be discretely (and distantly) arranged in the space if freight costs are high. Whatever the situation may be, it should be recognized that when production costs are relatively very substantial, freight *costs* by definition are insignificant. A spatial monopoly market, an all-sellers or all-buyers-at-a-point market or overlapping market area types tend to arise. The inquiry into market area size, shape, and equilibrium is then rendered trivial.

Fixed cost projections and other studies

It is further significant to suggest that the above-mentioned relations apply to cases of high fixed costs too. In particular, if future research does happen to uncover many instances where fixed costs are relatively substantial, so that extensive market areas are in fact required, a trivial verification of the possibility of nonhexagonal market areas may be at hand. Early findings indicate that fixed cost is extremely low with respect to Mills-Lav market area requirements. Correspondingly, a census study of freight costs, production costs, and value of shipments indicates that

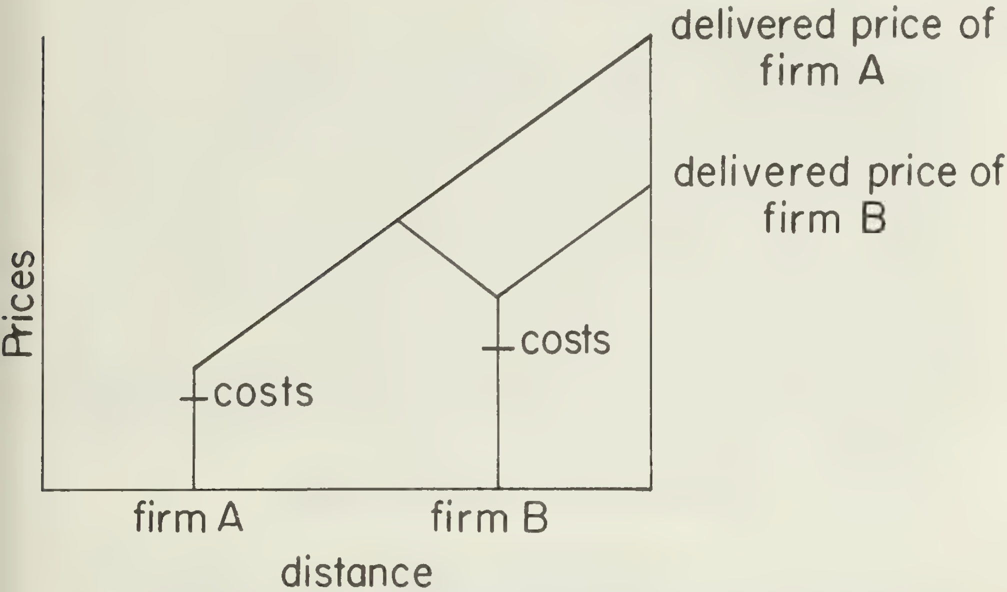


Fig. 10-4 Location away from the production center.

prevailing market areas in the American economy are of extremely small size compared to the profit-maximizing hexagon size of economic theory.

Bulky products. One other main possibility remains for analysis. It appears that there *are* cases of substantial freight costs (e.g., furnace slag). High freight costs also characterize the products of forests and mines. Is it likely that in these instances the firm's market area would be significantly large while, at the same time, the requirements of the nonhexagon competitive market area case are met? Or rather, do these possibilities suggest other relationships? In answer, note that high freight cost situations almost invariably apply to perishable or bulky products. They also tend to involve plant location at or close to raw material supply points, or at or close to the market. The location of rival firms may therefore well be at the same point as the subject seller. As in the cases of relatively high production costs or fixed costs, the matter of market area size and shape is therefore likely to be independent of the competitive forces discussed in spatial analysis.[11]

11. Buyers from firms subject to high freight costs are typically drawn towards the sites of suppliers in their choice of location. Such plant location enables them to keep the delivered price of their materials to competitive levels of cost. The full k possible in light of the relatively low c and high b values that must prevail in the subject case either enables firms to locate at a distance if raw materials are available there or is so high as to cause potential buyers at the natural peripheral points of the firm's market area to be buying at a cost which would make effective competition impossible.

Locations at distant places. By the very nature of the k/b ratios, K_0^* may run to very great values. However, where distance *is* a vital cost which requires economizing, where location at any or most points in the country is possible, where entry is open and competitive conditions prevail, a rival can be expected to locate at a distance from another firm, as in Fig. 10-4. When the new firm so locates at a distance, the market area of the initial firm in the industry is reduced substantially in size. It is thus our thesis, from both theoretical expectations and preliminary empirical findings, that market areas in the United States are much smaller than the spatial monopoly circle, the many-sided polygon sizes. Even the profit-maximizing hexagon size is much too large for the sizes which competitive economies will tend to produce. Exception would hold only when all sellers are, in effect, located together in economic space.

VI. *Preview of the final chapter*

There remain as requirements the tasks of determining (1) whether the competitive prices and distances set forth for the hexagon in Table 10-2 are stable, and (2) whether the values shown in Table 10-2 for competitive structures in economic space are similarly ordered when sellers practice spatial price discrimination. But before these problems are analyzed, two appendixes are affixed to this chapter. Appendix I explains the computer runs which provided our Table 10-2 data; Appendix II is included for readers interested in gaining a bird's-eye view of the productive efficiency which the present writers believe applies to the space economy. The analysis set forth in this Appendix II should be considered in the backlight of the distributive (locational) efficiency described in Appendix II to Chapter 8 above *and* the market area efficiency presented in Chapters 9-11.

Appendix I: *Computer runs*

The purpose of this appendix is to provide an explanation of the computer runs which established the values presented in Table 10-2.

Our problem there was to solve three simultaneous equations in the three unknowns, Y'_i, K'_{*i}, and p'_i. However, since $Y'_i = 0$ by equation (10-III-5), we can reduce our system to the following:

$$\alpha_{1i}K'_{*i}{}^4 + \alpha_{2i}K'_{*i}{}^3 + \alpha_{3i}K'_{*i}{}^2 - F = 0,$$
$$4\alpha_{1i}K'_{*i}{}^2 + 3\alpha_{2i}K'_{*i} + 2\alpha_{3i} = 0,$$

where

$$\alpha_{1t} = -11.2218, \quad \alpha_{2t} = 6(2.6672p'_t - \sqrt{3}b), \quad \alpha_{3t} = 3\sqrt{3}(b - p'_t)p'_t,$$

$$\alpha_{1s} = -3.6465, \quad \alpha_{2s} = 8\left(1.0294p'_s - \frac{b}{\sqrt{2}}\right), \quad \alpha_{3s} = 4(b - p'_s)p'_s,$$

$$\alpha_{1h} = -1.8101, \quad \alpha_{2h} = 12\left(0.4639p'_h - \frac{b}{3}\right), \quad \alpha_{3h} = 2\sqrt{3}(b - p'_h)p'_h,$$

$$\alpha_{1c} = -\frac{\pi}{3}, \quad \alpha_{2c} = \pi(1.3333p'_c - b), \quad \alpha_{3c} = \pi(b - p'_c)p'_c.$$

This equation system consists of two equations in two unknowns, namely K'_{*i} and p'_i. This system is nonlinear and the maximum number of solution sets is eight (4 × 2). However, the domain of the profit function under free entry must be so constrained that $b > p'_i > b/2$ and specifically $(\frac{3}{8})b > K'_{*t} > 0$; $3b/4\sqrt{2} > K'_{*s} > 0$; $3\sqrt{3}b/8 > K'_{*h} > 0$; and $3b/4 > K'_{*c} > 0$.

The upper limits for p'_i and K'_{*i} are set by the noncompetitive values (i.e., those obtained without free entry). The lower limit is the spaceless equilibrium values. Since free entry compresses both p'_i and K'_{*i}, they must lie within the domain specified above. Of course, any solution set falling outside of the domain specified above is irrelevant.

With these specifications and constraints, an "iterative" program was formulated. Our concern was to see whether or not a given initial (arbitrary) set of values p'_i and K'_{*i} would converge to any set within the domain specified above.

Starting from pairs of p_i and K_{*i} which maximize profits under no entry, i.e., from $(p_t = b, K_{*t} = 3b/8)$, $(p_s = b, K_{*s} = 3b/4\sqrt{2})$, $(p_h = b, K_{*h} = 3\sqrt{3}b/8)$ and $(p_c = b, K_{*c} = 3b/4)$, a very quick convergence was observed within the domain for the polygons, but not the circle. The convergence for the circle fell outside of the domain as, in effect, the computer skipped the desired solution and converged to an inapplicable one. By changing the initial pair of values for p_c and K_{*c} so that $p'_c = 0.899 \ldots b$, and $K'_{*c} = 0.6499 \ldots b$, the c values recorded in Table 10-2 were obtained. We also started from several alternative initial values for all cases (market shapes) and found that the convergence was unique within the domain.

It might finally be recalled that the solutions which were obtained assume a specific value of F, i.e., fixed cost. Theoretically, the lower F is, the smaller are both $p'_i{}^*$ and $K'_{*i}{}^*$, and vice versa. But most significant, the positive relationship between Y'_i and K'_{*i} remain unchanged given p'_i. Only the height, and not the shape, of the profit function is affected by F. Thus, if $F_1 > F_0$, the profit curve (given p'_i and given F_1) may be

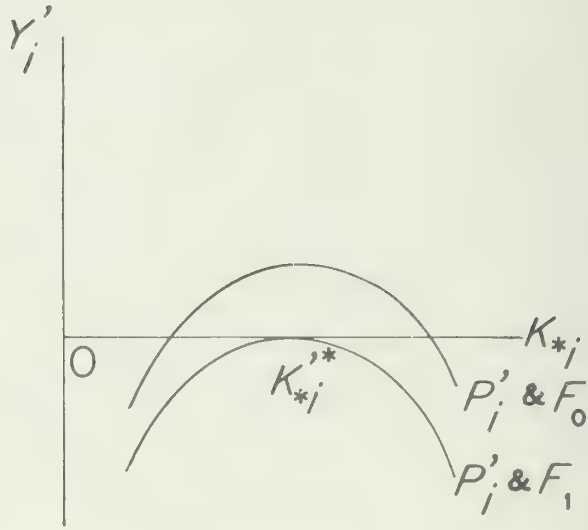


Fig. 10-5

distinguished from the profit curve under a p'_i subject to the other F, i.e., F_0, as in Fig. 10-5. Since the F_0 profit curve yields positive profit at $K'_{*i} = K'_{*i}{}^{*}$, greater entry would be required; hence both p'_i and K'_{*i} must be smaller. (The profit curve would shift leftward as well as downward, since p'_i is not a constant.) The more efficient the firm is, or the lower the fixed cost of producing a given product is and the greater the amount of competition in the industry is, *ceteris paribus*, the lower are prices, and the smaller is the size of the firm's market area. The same result applies with respect to the magnitude of variable costs, including the zero MC state. The basic *relations* shown in Table 10-2 apply to any of these cases.

Appendix II: *Economic welfare and market area shapes and sizes*

I. *Introduction*

We suggested in Chapter 1 that regional economics has many facets. Prominent among these—on the microeconomic level of thought—are the pricing patterns of spatial monopolists, spatial competitors, the loca-

tion (dispersion and concentration) of firms and of people, and the size and rank of their cities, among other subjects. The inquiry and conclusions of Chapter 10 point to a still additional concern: namely, the effectiveness of a competitive, free enterprise economy that is subject to the costs of distances. Significantly, the sizes and shapes of the market areas of firms are therefore but one side of the whole problem. Not only does final answer to the question of the effectiveness of a free enterprise economy depend on the degree of dispersion (or concentration) of firms in economic space (answer to which question was suggested in Appendix II to Chapter 8), and on market area sizes and shapes, but it depends on the pricing, technology, and costs levels of the firms in economic space.

Though we cannot go into all details of the firm in economic space in this text, selected parts of a very recent paper written by one of the authors [11] might be included here. This inclusion is designed to suggest our answer to this "related" side of the inquiry of our book on spatial pricing and market areas. As with the material presented in Appendix II to Chapter 8, we are limited here to the broad outlines of our overall theory. But even more important, though these same outlines were also written in the backlight conception of competitive firms using the f.o.b. mill price system, the conclusions presented will be seen to apply to our spatial price discrimination model as well. Indeed, the final Chapter 11 of this book will indicate the modifications one would have to keep in mind if it were a datum that spatial oligopolists always priced discriminatorily rather than f.o.b. mill. As with plant location theory, the theory of the firm in economic space [10] remains fundamentally the same, given our thesis that spatial competitors may discriminate in their price patterns. This similarity stands, even though in classical theory only monopolists are conceived to discriminate among consumers.

II. *Economic space and oligopoly*

Properties and facets of the spatial firm were formulated years ago by Weber, Lösch, and others.[12] They observed that firms tend to disperse over the landscape, especially when transport costs on the finished product are significant. In fact, apart from the obvious requirement of sufficient demand over an economic space, the critical supporting characteristic behind the establishing of spatially distinguished production centers is that a delivered-cost advantage must hold for these centers with respect to certain buying points. This amounts to saying that a fundamental "economic space" requirement for the "spatial" framework of thought, compared to the "nonspatial" classical framework, is that any

12. For details, see Greenhut [7, chaps. 5–8]. Also see [6, chaps. 2 and 3].

production cost disadvantage that *might* apply to a site distant from an existing production point must be offset by a transportation cost advantage which such a site possesses with respect to certain buyers. This net (delivered cost) advantage to certain buyers not only justifies the distant location, *ceteris paribus,* but indicates that competitive pricing by the distant firm will have significant impacts on the market area of the firm(s) at the production center. Identification of rivals and concern over their policies take place as a spatial oligopoly arises when the cost of transportation is significant.[13]

Given the necessary location condition of lower delivered cost to selected buyers, it follows that a competitive spatial oligopolist is defined as one who is willing to underprice other sellers with respect to these buyers. *Ipso facto,* an organized oligopoly (i.e. cartel form of oligopoly) over economic space involves an unwillingness of sellers to take ad-

13. We may sketch the origin of the spatial oligopoly rather easily by reference to Fig. 10-6. We assume in the figure that marginal production costs are constant, but higher at site B than at A; nevertheless, a delivered cost advantage prevails for site B vis-à-vis site A over the market area TT′ (cf. MCV with MC′V′ and MC′V″). Thus, if the firm(s) at A insists on cutting delivered prices below the P′D′ (and hence P′D″) schedule designated by the firm at B, a price war could ensue for the TT′ market; however, the war would be won by the firm at B, *ceteris paribus.* It follows that if the firm at B establishes some spatially lower prices, as in the figure below, rather than adopting the delivered price schedules of the firm(s) at site A (as it would under the organized oligopolistic basing-point price system), it could dominate the market area TT′. It is apparent (and particularized in the literature) that the impact on the price and sales of the firm(s) at site A typically is much greater if a new competitive firm locates at a distance from site A than if it locates at A.

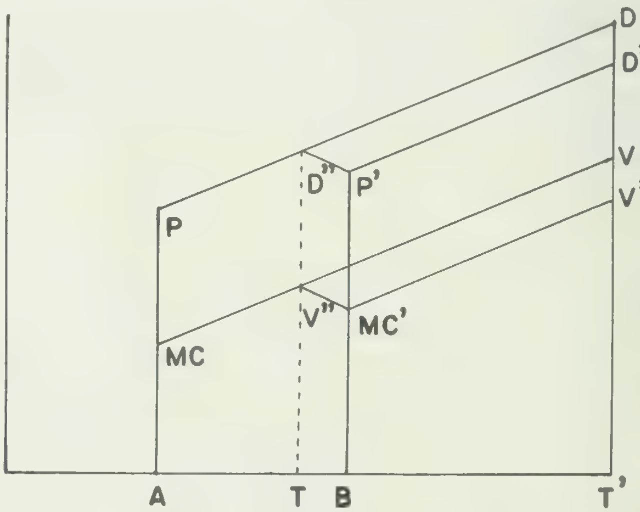


Fig. 10-6 Delivered prices and delivered costs.

vantage of the delivered-cost conditions necessary for effective location at the distant site.[14]

It is established accordingly in the literature that in order for transport cost to lead to the dispersion of firms, the coexistence of only a few generally prevalent conditions is needed. Thus (1) the demand for the product must not be point-formed in the space; (2) raw materials must be obtainable at several places in the space, or else (3) relatively high transport costs on the next-stage product must prevail; moreover, (4) labor and other production costs must be variable over the space *and* the site of lowest production costs must not be coterminous with the lowest transport-cost point to any and all requisite buying points. The expectation follows that at least 1, 2 or 3, and 4 ordinarily prevail together. Industrial dispersion and oligopolistic relations therefore mark the economy. But if oligopolistic industries are formed over the landscape, behavioral uncertainty among and between businessmen becomes a basic element in the system.[15] And an extra return, call it profits, is required for the firm in economic space.[16]

Uncertainty and economic equilibrium

It is specified in the theory that though diverse decision-making criteria are available to entrepreneurs who operate under uncertainty, the long-run survivors are those who evaluated the projected profit-loss matrix in a substantially similar way.[17] Common (similar) attitude and personality therefore characterize the members of an industry in the long run. And the returns for uncertainty among all firms and industries must be unique, except for a monotone transformation, provided one may properly assign greater disutility to greater uncertainty. It follows that profits for differential uncertainty are of the same order as rents for differential risks, skills, or fertilities. In turn, the profit commensurate with uncertainty is ascribable onto the classical average-cost curve of the firm. When the ascription is made, an adjusted average-cost curve is obtained which differs from the classical curve just as the classical

14. When the typically smaller total sales potential of the distant site is considered, it becomes even more apparent that location there is uneconomic if the site does not possess a delivered cost advantage over some part of the market. Indeed, this condition applies under either a spatially competitive or an organized oligopolistic price policy. See Greenhut [6] for details on effective location at distances from a production center under circular and other market area types, as well as other aspects of industrial location.

15. Knight [13, p. 318]; Machlup [15, pp. 255 ff.].

16. We use the term uncertainty in the Knightian sense [13, pp. 231 ff.] which, under our focus, is essentially an oligopolistic (behavioral) uncertainty. Unlike risk, objective probabilities do not apply to uncertainty.

17. Greenhut [8, pp. 274 ff.].

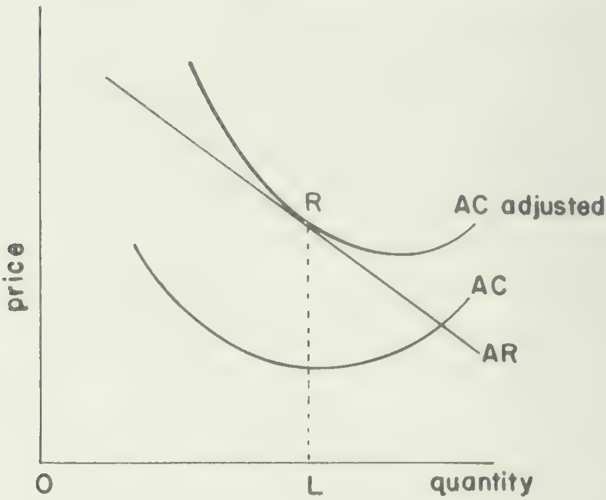


Fig. 10-7 Oligopolistic equilibrium – classical equilibrium.

explicit plus implicit average-cost curve differs from the curve which does not include all required imputations (cf. "AC adjusted" with AC in Fig. 10-7).

It is further specified in the theory that the negatively sloped average revenue (AR) curve of the spatial oligopolist is shifted leftward over time as a result of competitive entry. It is in fact demonstrated that in the long run the AR curve must become tangent to the adjusted average-cost curve (as in Fig. 10-7) at the output which yields lowest cost on the classical average-cost curve. If it were not tangent at the designated point, either the hired or the entrepreneurial factor returns would not be commensurate with lost opportunities, which is another way of saying that factor returns would not be conformable to uncertainty.[18] But a few more words in further review of the theory and the vital Fig. 10-7 must be entered.

The equilibrium properties of the firm in economic space

The concept of opportunity cost is critical to our theory of the firm in economic space. Consider, for example, the classical assumption that every factor tends to have some optimal lost opportunity *and* that if a differential rent also exists for the factor in the selected activity, this extra rent – as with the lost-opportunity rental – is a cost. Though not stressed in classical presentations of opportunity costs, it is evident that

18. See [10, chaps. 5, 8, 13, and 14] for details.

any income received in the lost alternative is attributable to some kind of work performance, i.e., energy expenditure, which yields the income. In other words, the opportunity-cost concept has two sides: the energy expenditure *and* the income which results from this expenditure. Because employment in one use often requires different energy applications than the expenditure intrinsic to the lost opportunity, and in fact may occasion a different economic lifetime (in calendar years) for the factor, a mapping from the best alternative to the subject activity is needed in order to establish the actual lost-income value for the subject activity. This mapping involves a transformation which includes the impact of uncertainty on either or both the energy expended and the income side of each activity.[19]

The income requirement established by the transformation serves as a fixed cost for the activity under consideration. The long-run viability of the factor – and more generally the firm – requires that this fixed cost be covered in full. There exists, accordingly, the adjusted average-cost curve (AC adjusted) which adds to classical average costs (i.e., AC) the

19. We may sketch via (i) below selected aspects of the transformation. Let superscript a designate the factor's best alternative. Assume an expected working lifetime M^a, net revenues r^a and energy units $\dot{e}^a$ used up in the appropriate production (or calendar) period time. (For simplicity, let energy units be the same in each unit production period of time.) Then discount the revenue and cost values by i_1 and i_2, where the discount parameters include not only the discount for time but a subjective addition for uncertainty such that if the uncertainty of income is great, i_1 will be great, while if the uncertainty in using the factor is great (i.e., the danger to the factor is great), i_2 will be small, *ceteris paribus*. After discounting, summing the quotients, and dividing, net revenues are obtained in the ratio form $\dot{R}^a$:

$$\dot{R}^a = \left(\sum_{j=1}^{M^a} [\dot{r}_j^a/(1+i_1)^j]/[\dot{e}_j^a/(1+i_2)^j]\right)/M^a. \tag{i}$$

Next, multiply $\dot{R}^a$ by the unit period energy expenditures applicable to the subject activity (e^s). This multiplication yields r^s in (ii), the linearized net revenue requirement in the subject activity:

$$r^s = (\dot{R}^a)\left[\sum_{j=1}^{M^s} e^s/(1+i'_2)^j\right]/M^s. \tag{ii}$$

Note that e^s stands for both the energy expenditure the owner would like to have the factor expend in any unit period of time and the expenditure optimally ordered by the prevailing technology. Our theory requires such identity, for if it did not prevail the factor would be high in cost and exit from the industry in the long run. Note further that M^s and i'_2 stand respectively for the working lifetime and the rate of discount on the energies expended in the subject activity.

It might finally be *stressed* that the revenue requirement r^s includes the return ordered for uncertainty. This return r^s is the factor's transformed opportunity cost, and it may or may not equal the factor's pay in the short run. Under our view of opportunity costs, utility maximization and profit maximization are practically indistinguishable, constituting basically one and the same principle, with leisure, danger, uncertainty, etc., serving as an intrinsic part of the decision process. We seek not to maximize uncertainty but to cover it.

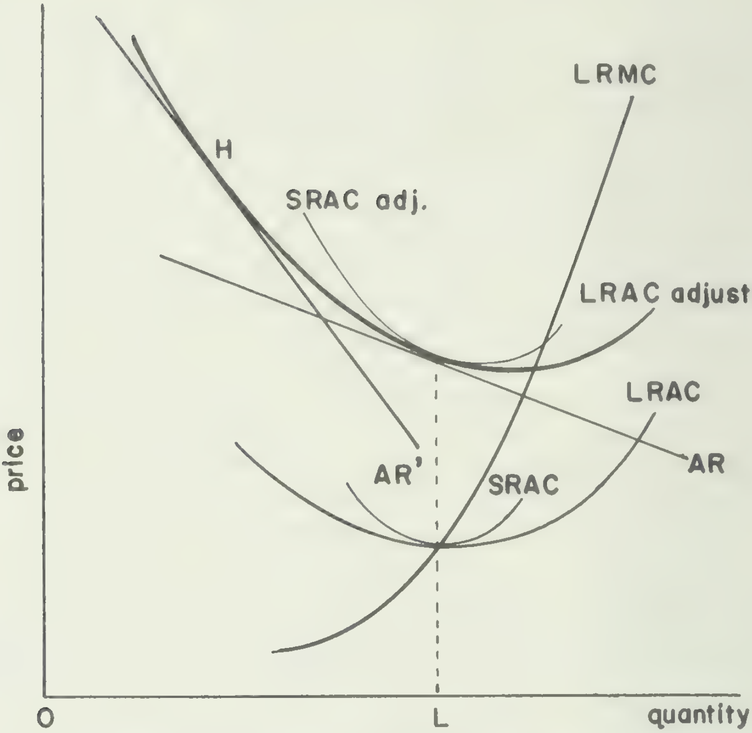


Fig. 10-8 Oligopolistic equilibrium – classical equilibrium (short- and long-run average cost curves).

transform of the uncertainty costs of the firm. But why does production take place at point R in Fig. 10-7 when the spatial oligopolists are competitive? Alternatively phrased, why does a tangency of AR to AC adjusted occur at the input-output point where classical average costs are minimized? We shall use Fig. 10-8 – which includes the long-run as well as short-run average-cost curves – to complete our outline of the theory.

For reasons advanced in detail elsewhere,[20] we define the entrepreneur solely in terms of capital inputs, relegating the management of a firm to the category of "hired" inputs. The extra cost of uncertainty or – let us now say – the rent ascribable for uncertainty on the capital investment thus constitutes an additional revenue requirement for the firm. By adding this rent to classical cost as a fixed cost, we find that the minimum point on the LRAC adjusted curve lies at a greater input-output point than that applicable to a firm in an economy in which certainty (or risk

20. Greenhut [10, chaps. 5, 8].

alone) prevails. Moreover, since uncertainty does not involve per se a functional performance, the payment required for it is a residual one. Though details must again be left for recording elsewhere, it should be sufficiently clear for our purposes that if the AR curve happens to lie below or above the LRAC adjusted curve, or is tangent to it at some point other than that related to the minimum classical cost point, such as H in Fig. 10-8, the income available to the entrepreneur on the capital investment would not be identically commensurate with the uncertainty and energies applicable to the subject activity. We can appreciate this condition because at input-output points less than OL, the capital energies used are less than those optimally ordered by the prevailing technology and entrepreneurial preferences; thus the tangency of AR to LRAC adjusted (which, to recall, covers the full opportunity costs throughout the curve) would entail a surplus that must be eliminated in time as a result of competition. In related manner, entry, relocation, or exit stem from nontangency situations, for these entail a windfall surplus or loss and/or involve a working lifetime (i.e., unit period energy application) which fails to conform to the factor's specification. Only at output OL will hired factors receive their lost-opportunity costs plus any relevant rent differentials while, at the same time, the entrepreneur obtains in full the return required for uncertainty in the oligopolistic-space economy.

III. *The number and size of firms required for technological efficiency in economic space*

Entry, relocation, or exit of plants and firms occupies a role in competitive space economic theory similar to entry and exit in competitive nonspatial economic theory. The ultimate tangency, however, of a negatively sloped AR curve to the adjusted LRAC curve makes the unorganized spatial oligopolistic equilibrium appear to be of vastly different order than the competitive solution. Notwithstanding appearances, the difference is fundamentally one involving the number of firms required for economic equilibrium. To appreciate this condition, recall that in the geometric series stemming from Cournot's analysis, new entry increases supply and causes price to descend to the competitive level. Indeed, in classical theory supply increases to the level where competitive firms produce $n/(n + 1)$ of the total market demand effective at the level of zero profits. In turn, price descends to the level of costs which yields zero profits.

Numbers in economic space

In any practical world of business, existing oligopolists will seek control of as large a share of the market as is possible. We might in fact conjecture that if the oligopolists failed to seek disproportionate market shares, some form of collusive – inefficient – oligopoly would exist. But, similarly, if firms did seek disproportionate shares, and unlike Cournot's firms actually gained unequal shares, the question arises whether inefficient large firms might not still survive in the long run? We must accordingly determine, first, whether an economy of competitive (i.e., unorganized) spatial oligopolists would cause unequal size firms to arise, and then, if so, how could the system guarantee survival of only the efficient firm?[21]

Firms of unequal size in economic space

Recall that Cournot's assumption of constancy of rival supply eventuated in the conclusion that equal shares would prevail among the firms. Moreover, the nonspatial character of the theory readily enables one to apply the same cost and demand conditions to new entrants as to old, changed of course from the original to the extent that the new entrants cause such change. Unfortunately for easy methodology, an analysis of economic space cannot utilize this convenient assumption, because the localizing of some firms promotes heterogeneities, and behavioral uncertainties must arise in economic space apart from any uneven distribution of Mother Nature's gifts. Thus if heterogeneities and uncertainties arise (and exist) over the space, the firms may well be different in size. Will, then, inefficient large firms be squeezed out of the market, or might they not continue indefinitely in the market?

In answer, note that new firms may locate either at a nearby point or at some point distant from a given seller. This entry means that the preexisting market area sizes and shapes will tend to be altered, and in turn that the sizes of old firms and plants will tend to change over time. Since the patterns of entry over the entire space may differ, and since writings on location theory from Lösch onward have noted that uneven-size towns and communities must arise over space, some firms, we propose, will tend to be large and will "own" many plants while others in different market areas, where different cost and demand conditions prevail, will have smaller and/or fewer plants. The LRAC curves of all firms – exclusive of rent – may well differ, ranging from acceptable highs to the lowest possible low. But the result is not a state of inefficiency of many firms compared to the efficiency of a few firms or just one. Let us

21. How oligopolists may seek to be large and yet not prevent entry is an irrelevant question for us in a study which assumes an unorganized oligopoly. Among other aspects, such a question involves legal matters. See Greenhut [10, chap. 15].

say that different costs and demands among market areas require some variation in size of "most efficient" firms over the landscape. Indeed, if the firm's costs are higher in one area than elsewhere and all other things appear equal at a given moment, its *profits* relative to its uncertainties must be insufficient and it will drop out of the market in time. Either this happens, or else demand or entrepreneurial differences prevail which serve as the offset to the differences in costs. We can appreciate the rationale behind our answer by considering the following situation.

Suppose there is room for only one more firm of smaller size within a market area, and that costs are higher for that small firm than for the larger firms in the same market area. But if the entry of a higher-cost (small) firm is justified by the prevailing demand and cost condition, so that the firm not only may enter the market but will survive the ravages of fair competition, a rental must exist for the more optimal firm(s). This rental, however, is but a classical rental, a differential return. What governs entry must be the sufficiency of returns vis-à-vis uncertainty, notwithstanding different rents for different skills, or rents for the early discovery of an opportunity, or rents for fertility (location), organization, etc. Not all entrepreneurs (managers) require and desire the same return. Firms need not be identical in size and form, nor need the LRAC curves of all firms be of the optimal-optimorum variety. Only the best *possible* set of LRAC's over the landscape need arise for any given industry. In turn, then, a representative LRAC is identifiable *after all required returns are* ascribed.[22]

Within and between market areas, some firms will thus be small and possessed of one or a few plants at best, and some will be large.[23] Indeed, some firms will be more efficient than others, in the sense of market area and/or firm costs. When costs are high in a given market area, or when the firm's costs are high because it has only a single-specialty plant and—in the industry—multiple plant units produce goods most economically, the high-cost firms will survive if and only if the demand for their products is sufficient to provide a return commensurate with uncertainty. Small firms will be found over the space, provided that the

22. It is conceivable that a small single-plant firm may coexist alongside a large multiple-plant firm. But any disadvantage in cost suffered by the small single-plant firm would—in the representative firm picture of the industry—be offset by the differential rental cost extra that is added to the cost of management of the large firm. In the limit, the costs of all surviving firms in the industry (i.e., those meeting their opportunity costs) would be the same. Any remaining positive balance of revenues over total costs (including all differential rents) would have to be in direct relation to the uncertainty of economic activity in the subject industry. The surviving firms must be efficient in location and cost.

23. Witness the furniture and paint industries in the United States, where large multi-plant manufacturers produce standardized products at a low unit cost and sell at a low price, while small specializing firms, essentially of a single-plant form, produce products of regional design, typically at high cost and price.

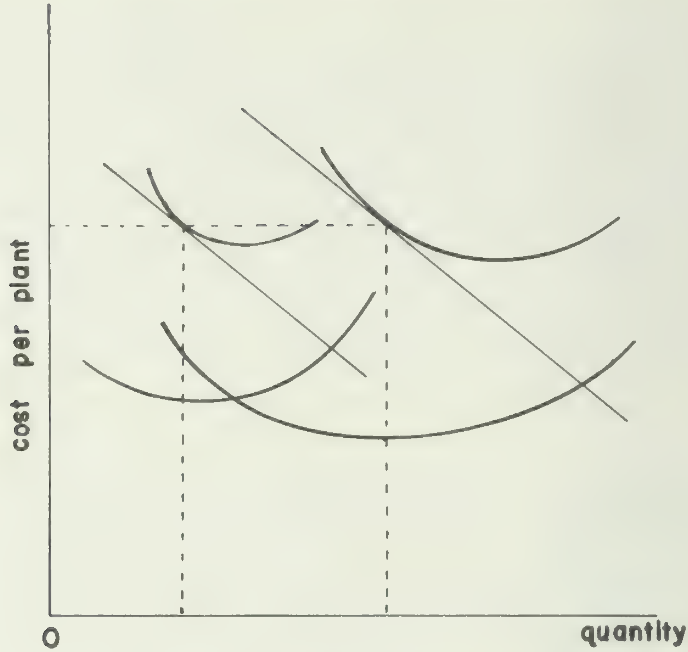


Fig. 10-9 Alternative sizes – varying opportunity costs.

managers and owners who could move into the market cover their own lost opportunities (including all uncertainties) in the long run. In turn, the small firms must squeeze out still smaller, less efficient firms in the short run, leaving the most efficient set of firms possible in the market (for example, see Fig. 10-9). How many competing firms must there be in the space? Our answer should be apparent: we require the number necessary to bring about results such as those of Fig. 10-9; it may be two, three, or many firms.

IV. *Pure monopoly in a space economy vis-à-vis two competitive firms or more*

We have assumed to this point in Appendix II the state of oligopolistic competition over space. This competition means to us a market in which entry and exit are open (i.e., unconstrained by artificial barriers that work against efficiency).[24] We have, accordingly, conceived of a system

24. Observe that whereas a patent would seemingly fall into the category of an artificial barrier, we would not consider it to be a force which unduly restrains entry if the years during which the patent right is granted do not extend beyond those necessary to promote the research. In contrast, we shall regard as abuses those patent extensions and pools which restrain entry beyond the limits required to encourage research and invention.

where a firm *may* have many plants, where its plants *may* serve different market areas, where many firms *may* exist, and where competition *may* take place between plants in a given market. In many and even all situations, the competition from plants of *rival firms* is the force which shrinks the firm's demand curve so that earnings commensurate with uncertainty result. This competition compresses the magnitude of demand for the firm's product(s), as to both its horizontal and its vertical extent.

Although of tangential interest, there is advantage here in comparing the natural-monopoly market type with that of oligopolistic competition. This is so because the theoretical results obtained from such comparison are closely related to our general theme. They help point out certain features of our general theory more sharply than may otherwise be done.

Advanced technology and limited demand initially

The question may fairly be asked whether it is possible for one firm to establish a plant in one market area under, say, a protective patent and subsequently to continue to anticipate all future developing new market areas to the point where the firm attains giant size and protected monopoly position just from this growth. Suppose the plants of the firm earn monopoly profits in each of its market areas. Would not other firms and plants enter some markets so that a competitive spatial oligopoly becomes possible? Or could a spatial monopolist alone prevail simply because demand and cost conditions are such that a large, multiple-plant firm appears to preclude or in fact precludes the entry of others?

To answer the questions, consider for example the Alcoa company in the United States. Did not its size alone constrain entry? Did not the mere existence of a large (multiple-plant) firm intimidate others with the effect of restraining entry, notwithstanding potential economic surpluses for new firms? Our answer to this last query is based on two possibilities.

One, suppose the existing firm is "too large"; that is, it is operating too many plants and suffers, accordingly, from an unwieldy organization. Clearly, demand relative to cost is such that new firms and plants may readily enter the market. Spatial monopoly would be inefficient, and in order for the firm to survive, it would have to employ predatory practices. Effective antitrust laws and enforcement would then be necessary.

Two, suppose the opposite; that is, a firm must be a giant to operate successfully. Any single-plant firm or a firm with only a few plants is small and, by assumption, inefficient. It would be unable to gain the substantial advantages derived from a vast enough order of research and from centralized buying and selling (i.e., internal economies); moreover, it would be unable to schedule production properly, to shift manpower, raw materials, inventories, top personnel (sales and otherwise) between

a large enough group of plants to assure steady production and maximum efficiency. In effect, its long-run average costs *in each market area* would be too high and the firm could not earn enough to cover the applicable uncertainty. By assumption, only the larger multiple-plant firm is flexible enough to survive; thus, dissolution of the large firm would be uneconomic. It follows that if demand relative to technology is sufficiently limited, a natural monopoly must be conceived as a possibility under a spatial framework of thought just as one may assume its existence in nonspatial theory.[25] However, though we suggest that a spatial monopoly *may* exist, with prices and profits probably too high, we propose for the spatial economy the existence of a strong likelihood that new plants and firms will be able to locate somewhere in the space, with the effect of tending to bring about a spatial oligopoly.[26]

Natural monopoly must be regulated, but two firms or even one firm could establish market efficiency

The alternative situations and solutions that relate to the matter of efficiency and size, to repeat, are these. (1) If large size is uneconomic, small firms enter if entry is kept open. (2) If large size is economic, then either a monopoly or duopoly or perhaps even oligopoly *without excess profit* is possible.[27] If a natural (excess profit) monopoly prevails, classical

25. In a broad Löschian sense, the landscape may be unevenly endowed and the demand over space may be quite discontinuous, *or* technology may require such large plants that nonoverlapping (circular) market areas will prevail within which only one plant (one firm) could survive; alternatively, demand and cost conditions may be such over a continuous space that only one firm (with one plant or multiple plants) may survive. In any of these cases, regulation of prices and profits would be in order.

26. Few instances of *natural monopoly* can be expected to prevail in the long run in a spatial economy where *expanding demand and location at a distance are likely conditions.* Governmental policies should be to promote entry, for example by patent licensing or requiring the original firm to impart productive know-how, and preventing the original firm from coercing the new firm(s) into adopting an organized oligopolistic price and location policy. See Greenhut [10, chap. 7], where it is shown that, in general, organized oligopoly is less profitable to new (small) firms which locate at a distance from an established production center, notwithstanding the higher prices charged to consumers.

27. When entry reduces oligopolistic profits to the nonexcess profit amounts (or even the monopoly fails to provide surplus returns) the market is in effect described as a *competitive* oligopoly. Manifestly, except for natural monopoly cases, even a monopolist *must* conceive of potential rivals who may possess similar and/or even slightly less efficient technology (hence LRAC curves) *provided* entry is open. These rivals can, if they will, quote a little lower price than any price which the existing monopolist may quote in excess of the normal price (i.e., the price commensurate with his lowest cost on his lowest LRAC curve). The possibility that some monopolist may expect competitive entry if he earns excess profit could prompt him to price at such a low level as to exclude others, but yet to gain sufficient returns to cover his lost opportunities. This possibility thus provides an instance of a single firm only covering its opportunity costs. Of course, one other single-firm possibility of similar order can be mentioned – namely, the firm that just happens to find its most profitable (yet only a normal profit) position to be commensurate with its lowest-cost production point. At this cost-output point, the selling price mark-up just covers all lost opportunities.

attitudes and beliefs toward it may be applied.[28] If one or more large firms can enter, encouragement may be provided by compulsory licensing of new firms and related practices. Firms will continue to enter so long as the free-entry condition is maintained and profits are excessive compared to prevailing uncertainties.[29] Not only have we, therefore, answered the query as to the position of monopoly in the economy, but the number of firms and plants necessary for the equilibrium has also been fully identified. Given free competitive action, firms will enter or exit towards the end of establishing the hierarchy of returns that is commensurate with the hazards of doing business in the economy (see Section II above). Typically, two firms or more will fulfill this goal, with each firm establishing the number of plants necessary to minimize its costs over the space. Returns allocable to all firms in each market area would, in such a case, be compressed to the level that is commensurate with all opportunity costs (uncertainties included).

Maximizing consumer satisfactions: an aspect of numbers

It is perhaps trivial to note that the criterion of technological efficiency and productive compliance with consumer wants has generally been accepted in free enterprise economies as a culturally approved welfare criterion. In substance, economists have ignored the subjective basis behind such a judgment or else accepted the results as desired by all. In either case, normative status is often ascribed to any market type which yields productive efficiency and provides maximum consumer satisfactions. Since we have already derived the efficiency result of spatial competition, only the other condition—that consumer wants are satisfied perfectly—remains to be analyzed before we can claim effective operation for an unorganized, open entry and exit oligopoly over an economic space. Consider the following.

It has often been said that if $P > MC$ somewhere in the economy and $P = MC$ somewhere else in the economy, production *will not conform* to consumer wants. Apparently, mixing pure competition with oligopoly, and with monopoly, etc., yields imperfect compliance with consumer wants. But recall that in classical theory, where self-owned more-fertile land is used rather than less-fertile land, we would have $P > AC = MC$ before the imputation process. After differential land rents are ascribed

28. The overall task of antitrust is therefore evident, we suggest. It is to make sure that neither unregulated high-profit natural monopoly nor organized collusive oligopoly exists, and to condemn practices which artificially restrain entry. If the market is that of the natural monopolist, it must be regulated toward the end of yielding returns commensurate with competitive uncertainties elsewhere in the economy.

29. If entry *is* kept open while collusion or organization in any form is ruled out, the state of oligopolistic competition in economic space that was previously described will arise.

to the supramarginal producer, P = AC = MC. Correspondingly, in classical economics P > MC may hold for one plant (firm) while P = MC holds for another, not because of different land fertilities, but because of different skills or risks. Here too, when differential rents for skills and risk are ascribed, the classical trilogy P = AC = MC appears. The same kind of imputation process applies to a space economy, for if profits are commensurate with uncertainty, they may be regarded as a cost in the same sense as differential rents.[30] At the input point where opportunity costs are perfectly covered, differences in the P/MC ratios of firms in one sector (or market area) of the economy before the imputation of these profits signifies equality throughout the economy after the imputation is made. By recognizing uncertainty as an opportunity cost that must be defrayed over time, a mapping from the classical position to oligopoly over an economic space results. Differentials between P and MC in any market which relate strictly to profits for uncertainty simply involve differentials of the same order as those which mark a nonspatial system before rental imputations are made. The theorem follows that the spatial economy moves inexorably to the position of being in perfect compliance with consumer wants, again provided the spatial oligopoly is unorganized and competitive entry and exit are maintained.[31]

V. *A basic characteristic of the spatial model: sticky prices, but competitive results*

Let us note a basic essential of our model that was earlier suggested to hold true—our inference that though firms in an oligopolistic system are able to identify their rivals, and hence act in deference to the reactions of rivals, continued entry over time has the effect of causing them sooner or later to act *as if* they were perfect competitors.

30. We are implying that the opportunity cost for uncertainty may be treated as a variable cost besides the fixed-cost approach used in Figs. 10-7 and 10-8. Then, not only would AC be adjusted but also MC. Indeed, it is shown (Greenhut [10, app. to chap. 5] that final equilibrium results are the same, albeit the diagrams showing the imputation of uncertainty as a variable cost are not as revealing as the fixed-cost imputation diagram. For present purposes, we might add the observation that under the variable cost imputation *the ascription to* AC *and* MC *at the input point which yields the classical lowest cost point* —i.e., where all factors (including the competitive entrepreneur) cover their classical opportunity costs—*raises* AC *and* MC *by the same amount.* Then P = AC adjusted = MC adjusted, provided, to repeat, returns are commensurate with uncertainty.

31. We need not concern ourselves with charges of tautology when we treat the profit sum commensurate with uncertainty "as a cost." The nontautological condition stems from the fact that our conceptual scheme allows for the states of short-run windfalls (profits or losses), these in effect being added to or taken from the returns that are commensurate with uncertainty.

Price changes in nonspatial economic theory

Recall that in Cournot's framework the entry of a second, third, and later firm led to increased supply and reduced prices. Of course, nonspatial oligopoly theory does not always require each entry to have this impact. It is, in fact, quite conceivable that some conscious parallelism of action may develop among oligopolists, with total output and price remaining at the previous level notwithstanding the new entry. This type of relationship amounts, however, to collusion between the firms, a behavior pattern that could be ruled out by the antitrust laws of the state. If it is ruled out, sporadic price reductions and some increase in output may be expected over time due to the entry of firms. We shall see in a moment that this expectation of sporadic price reductions holds even more emphatically for our space economy than for a nonspatial competitive oligopoly.

Price changes in spatial economic theory

We observed earlier that under an unorganized, spatially competitive f.o.b. mill pricing system, the impact on sales of a new firm at a distance is greater than that of a new entrant in a nonspatial system. This vital principle, already noted in Section II of this appendix, and even earlier in Chapter 7 for the purpose of indicating why economic space culminates in an oligopolistic economy, also explains why a competitive but not hypercompetitive oligopoly tends to arise. It may be viewed sufficiently well for the purposes of this appendix as follows.

Consider the situation of negatively sloped linear demand with two (or more) firms located at the established production point. Let a third (or later) seller enter the market either at that point or at some distant point. In the initial instance, *ceteris paribus,* the entire market is shared by all firms. In the latter instance, which under the assumptions of this appendix ultimately occurs, the distant market, *ceteris paribus,* is taken over more or less in full. The monopoly-like advantage of the distant entrant over the portion of the market area nearest to his location has a very great impact on the sales of the firm(s) at the old center. It causes the seller at the original center to view distant entry as shifting his demand curve *noticeably* to the left.

It is easily shown that under the above conditions the impact on the total sales of the firms at the original production center is vastly greater if a new firm locates at a distance than if just one more firm should locate at the old production market point.[32] With repeated entry at locations

32. Greenhut [7, pp. 115 ff.].

distant from and in alternative directions from the production center, any seller in the established production center must sooner or later view his average revenue curve as shifting leftward along both axes, hence downward. Output rises then, and price must fall, according to the classical rules of economic analysis.

In the *point-formed* nonspatial economy of classical theory, a monopolistic leftward but not downward shifting of the curve may continue to be postulated as entry mounts, for the sales impact of any single new entrant may well be slight. Even here, however, there is good theory to indicate that our model of rising output and falling price ultimately must hold with enough entry.[33]

Spatial oligopolists tend over time to learn to ignore the reactions of their competitors. More generally, they contemplate and expect, in time, only a moderate reaction from their rivals. They reject hypercompetitive reaction because this leads to disaster in a spatial economy. In turn, long-run inflexibility would, as a possibility, be based upon what would then have to be an everlasting kinked demand curve notwithstanding increasing competition. We reject, however, the "tacit collusion" implied by the kinked demand curve because it violates the assumed conditions of free entry and competition in economic space.

Our theory, therefore, is that if new firms enter and locate next door to each other, total output *may* rise with price falling. But if new rivals also locate elsewhere and form other production points, the combination of local and distant competition raises the likelihood of some increase in total output ultimately taking place. In turn, a decrease in price must occur.[34] So it goes – with entry open and collusion ruled out, price should fall at least sporadically, and economic returns for entrepreneurship must descend toward the level commensurate with the uncertainties of engaging in economic activity.

VI. *The mixed economy*

Our previous discussions which claimed that oligopoly is coterminous with economic space implicitly assumed that freight costs were significant and that sellers and buyers are distributed over the landscape; alternatively, demand had to be relatively limited vis-à-vis technology in any given market area so that only a few firms (conscious of one another's policies) could be found together in the market. In the real world, however, we might in effect find all buyers and/or sellers at a

33. See Grossack [12, pp. 406 ff.].

34. We are asserting here that our theory is most applicable to a world of economic space, more so than a point-formed oligopolistic economy.

point, or insignificant freight costs may prevail; indeed, rather simple technology may characterize the industry. Then pure or monopolistic competition could be possible. Assuming that such simple market types existed, the economy would be a mixed one. Would it be efficient? More inclusively phrased, are all market types possible even though economic activities and people are distributed over economic space; and if they are, would the system be efficient?

To answer, let us observe immediately that if pure competition exists, knowledge is complete and only risk, not uncertainty, may apply. The classical costs are all that will limit output, as factor returns are monotonically unique *up to a linear transformation.* But this is counterpart to the production policy and practice of unorganized oligopolists. In a nutshell, the same signals are received in the unorganized oligopoly world as those prevailing under pure competition. The one market is isomorphic to the other.

Monopolistic competition, in contrast, has an equilibrium state in which arbitrary discrepancies arise in pricing and productive inefficiencies prevail. Most significantly, this market type violates the fundamental rule (and meaning) of competition in space, as the opportunity-cost concept implicitly requires returns throughout the economy to be commensurate with each other while the tangency of AR to AC in the monopolistically competitive equilibrium carries no such necessary relationship. To understand this vital thought, let us repeat the statement that the opportunity-cost concept requires identity between the desired use of a factor and the industry's technology; otherwise the factor would prove high in cost and exit from the industry in the long run. But this identity involves the classical lowest AC point, while the monopolistically competitive equilibrium is traditionally viewed to offer tangency of a negatively sloped AR curve to an AC curve at some other point.[35] By recognizing that this AC curve must still have a fixed-cost imputation for lost alternatives, we see that the factor could receive a return commensurate with its best alternative at a smaller energy expenditure level than that which it is willing to incur. A surplus exists which from one industry to another would be arbitrary; such a surplus would have to be squeezed out, with conformable returns resulting

35. The *instability* of the Chamberlinean tangency [3, chap. 5] has also been pointed out by D. Dewey [5]. In fact, as W. D. Maxwell [16, p. 108] correctly observes, a major portion of [5] is occupied by this discussion. The discussion, however, is quite different from Greenhut's [10] since the criticism of [10] is based on tangency solutions resulting from an assumed competitive entry of new firms, while Dewey is contending that the *collusive* merger of existing firms would serve as the main force behind more efficient production. Dewey's argument seems to run counter to the myth that the greater the number of firms in an industry, the more efficient is production. In our view, he appears to be mistaken both in theory and in practice when he implies that merger leads to more efficient production. See Section V of this appendix, and in particular note 27, above.

throughout the system, before efficiency could exist.[36] In other words, the Chamberlinean tangency just will not survive. So, what on the level of assumptions could serve as the basis for inefficient monopolistic *competition* reappears in the real world as an effective market type! Indeed, the economy one finds *in economic space* must be of the order of pure competition mixed with oligopoly, as any monopolistically competitive industry would be reshaped in time by entry, relocation, and exit into the purely competitive or oligopolistically competitive mold.

VII. *Conclusion*

We have suggested above that a competitive spatial oligopoly economy produces at lowest possible cost and charges the lowest possible mill price(s). Elsewhere [10, chaps. 11–13] we have gone into details, including problems of the impact of "wrong (locational) start – wrong city development?" and their like. For the purposes of our present inquiry and design, all that need (and can) be added is the suggestion that if the viable firms in a competitive spatial oligopoly are counterpart to the firms of pure competition theory, the firms which are inefficient in the sense of failing to earn their opportunity costs must fail to survive. It follows, in turn, that firms whose locations fail to conform to consumer demand or to the prevailing technology, and thus to the lowest possible level of production and distribution costs, must fail to survive to the long run.[37] The viable spatial firms, we suggest, are firms whose distribution pattern, production technology, and use of resources are classically economic.

References

[1] M. A. Adelman. *A & P! A Study in Price-Cost Behavior and Public Policies*. Cambridge, Mass., 1959.

36. It might be observed that if the opportunity cost imputation to the classical AC curve is treated as a variable input rather than as a fixed input, tangency of the negatively sloping AR to the variable-cost adjusted AC curve will involve a smaller return for the factor than that which its best alternative offers. Unlike the tangency described in the text above, the factor will be receiving an inadequate (rather than surplus) return, given a tangency to an AC curve that includes increased opportunity cost as inputs increase. In any case, final results will be the same, although tangency of AR would exist only with respect to the fixed-cost AC adjusted curve. The "same final results" stem from (1) the entry of firms in the case where tangency of AR initially occurs with the fixed-cost adjusted AC curve at a less than optimal classical output, and from (2) the exit of firms in the case where initial tangency of AR to variable-cost adjusted AC prevailed. The entry and exit bring the economy to its equilibrium position.

37. Elsewhere [10, chaps. 12 and 13] we probe directly into this question rather than proposing indirect proof of conclusions, as has been provided above.

[2] M. Beckmann. *Location Theory.* New York, 1968.
[3] E. H. Chamberlin. *The Theory of Monopolistic Competition.* 5th edition. Cambridge, Mass., 1946.
[4] M. R. Collins and L. E. Preston. *Concentration and Price-Cost Margins in Manufacturing Industries.* Berkeley, 1968.
[5] Denike and Parr. "Production in Space, Spatial Competition, and Restrictive Entry." *Journal of Regional Science,* April 1970, 10:54.
[6] D. Dewey. *The Theory of Imperfect Competition: A Radical Reconstruction.* New York, 1969.
[7] M. L. Greenhut. *Plant Location in Theory and in Practice.* Chapel Hill, N.C., 1956.
[8] M. L. Greenhut. *Microeconomics and the Space Economy.* Chicago, 1963.
[9] M. L. Greenhut. "The Decision Process and Entrepreneurial Returns." *Manchester School of Economic and Social Studies,* 1966, 34:247–67.
[10] M. L. Greenhut. *A Theory of the Firm in Economic Space,* 2d printing. Austin, 1974.
[11] M. L. Greenhut. "The Theory of the Spatial and Nonspatial Firm." *Weltwirtschaftliches Archiv,* 1970, pp. 87–113.
[12] I. M. Grossack. "Duopoly Defensive Strategies and the Kinked Demand Curve." *Southern Economic Journal,* 1965/66, 32:406–16.
[13] F. H. Knight. *Risk, Uncertainty, and Profit.* New York, 1957.
[14] W. Leontief. "The Theory of Limited and Unlimited Discrimination." *Quarterly Journal of Economics,* 1939, 54:490–501.
[15] F. Machlup. *Economics of Sellers' Competition, Model Analysis of Sellers' Conduct.* Baltimore, 1952.
[16] W. D. Maxwell. "Imperfect Competition Then and Now." *Mississippi Valley Journal of Business and Economics,* 1970, 6:105–19.
[17] E. S. Mills and M. R. Lav. "A Model of Market Areas with Free Entry." *Journal of Political Economy,* 1964, 72:278–88.

11. Equilibrium shapes and sizes under nondiscriminatory and discriminatory competitive pricing in economic space

I. *Introduction*

Hitherto the market shapes have been treated as givens. But they may well show change in the long run due to a change in costs or, for example, the addition of new transport routes. Moreover, the results obtained in Chapter 10, in particular the conclusion that profits are greater for market shapes closer (and closer) to the circular form, appear to imply seller advantage in having circular market areas. This implication obviously holds true for spatial monopoly. It was thought also by Mills and Lav [6] to be applicable to many instances of competition. Chapter 10 established the contrasting thesis that only the hexagon is typically viable under conditions of competition. To see this more clearly, the results derived from equations (III) in that chapter, as recorded earlier in Table 10-2, will be reexamined herein. Present emphasis, however, will be on explaining how the hexagonal form comes about. After this is done, the problem of competitive market area stability can be evaluated, and finally the question resolved of the impact of discriminatory pricing on market area sizes and shapes.

II. *A restatement of results of the spatial competition model*

At first glance the mill price $m'_c = 0.1789b$, the circle size $K'_{*c} = 0.4008b$, and the maximum delivered price p'_c equal to $0.5797b$ could all appear as the final equilibrium values in the competitive space economy. But, as was claimed previously, this is not the case because profitable empty spaces beyond the distance $0.4008b$ would exist to which sales could be made at prices greater than the delivered price $0.5797b$. The firm, in other words, must be conceived to be able to sell over market segments falling outside the circle of size $K'_{*c} = 0.4008b$. So, the circle of size $K'_{*c} = 0.4008b$ would be replaced by a hexagonal market area of size $K'_{*h} = 0.4008b$. However, $K'_{*h} = 0.4008b$ and mill price $m'_h = m'_c =$

$0.1789b$ provide positive profits. Any (and all) seller(s) would sell to peripheral points at the limit-delivered price $p'_h = 0.6418b$.[1] But this price is not the equilibrium delivered price because it involves $\partial Y_h/\partial K_{*h} < 0$ at $K'_{*h} = 0.4008b$. Any seller would be still better off by reducing his market area size, and charging a higher mill price m'_h. Of course, the effect of the original state of positive profits combined with the increase in profits from the smaller size/higher price policy induces entry. A competitive, stable equilibrium must ultimately be reached. As derived in computer runs of equation system (III), price $m'_h = 0.1873b$ and $K'_{*h} = 0.3744b$ appear to provide this equilibrium. [See explanation in the note to Table 10-2 for other details.]

Instability resulting from $m'_c \leq m'_h$ in the zero-profit state of Table 10-2

The question arises whether the hexagonal market indicated in Table 10-2 as the final position is in fact a stable equilibrium position. This (possibly surprising) query stems from the thought that after the seller has converted his circular market to the hexagonal form, and then witnessed a shrinkage in the area because of entry, his highest delivered price $p'_h = 0.6234b$ reflects a higher mill price (rather than a lower mill price) than the old circle mill price – viz., $m'_h = 0.1873b > m'_c = 0.1789b$. Even after the entry of competitors and the elimination of windfalls, it might, therefore, appear that some prospective entrant could conjecture that an opportunity exists to locate next door to the hexagonal seller's site and to sell over a circular market area shape at a mill price less than $0.1873b$. He could then compete successfully with the hexagonal seller for buyers in the market. In fact, the new entrant who sells over the circle market at a lower mill price than the original seller could expect to eliminate the older firm, *ceteris paribus*. The question follows, of course: Can he also compete with the six other hexagonal sellers surrounding him so that at the worst he is covering his opportunity costs in full?

The price war and market area distortion

Suppose along the lines of [3] that the new entrant considers himself capable of winning any price war with the older seller who is trading over the hexagon-shaped market area of size $K'_{*h} = 0.3744b$ at price $m' = 0.1873b$. The new (circle) seller sets his mill price at its minimum level, i.e., $m'_c = 0.1789b$. *Given the related market size $K'_{*c} = 0.4008b$*, his delivered price to the shortest-distant boundary point must be less than any distant rival's delivered price to that point. Moreover, to sell over a

1. p'_h is easily derived from the relation $K_{0h} = 2K_{*h}/\sqrt{3}$ which applies to the hexagon, where K_{0h} stands for the farthest distance in the hexagon.

circle, his price $m'_c + K'_{*c} \leq m'_h + 2K'_{*h} - K'_{*c}$ – a condition, however, which does not hold.[2] The appearance of successful competition as viewed at the hexagonal boundary point is thus misleading. The circle, we have noted, cannot prevail, as the circle seller cannot sell over the full required space nor, accordingly, can he cover his opportunity costs. His *kamikaze* entry indeed hurts all, including the hexagonal sellers who surround the newly entered firm, since the delivered prices of these sellers to their nearest original market boundary points, i.e., their prices $p'_h = m'_h + K'_{*h}$, are higher than the *kamikaze* price $m'_c + K'_{*h}$ to such points. None of the firms is covering lost opportunities. Instability or indeterminacy appears to arise when the competitive hexagonal mill

2. This relation is derived as follows. In Fig. 11-1, AB = 2AC = 2CB = $2K_{*h}$, AD = K_{*c}, BD = $2K_{*h} - K_{*c}$. At D, A's delivered price is $m_c + K_{*c}$, while B's delivered price is $m_h + 2K_{*h} - K_{*c}$. This gives us the required relation set forth in the text above: namely, $m_c + K_{*c} \leq m_h + 2K_{*h} - K_{*c}$ where seller A becomes Seller One and seller B is Seller Two. However, $m_c + K_{*c} = 0.5797b$ while $m_h + 2K_{*h} - K_{*c} = 0.5353b$. The required relation does not hold, and hence Seller One cannot sell over a circle. On the other hand, it can also be shown that B's delivered price is higher than A's at C.

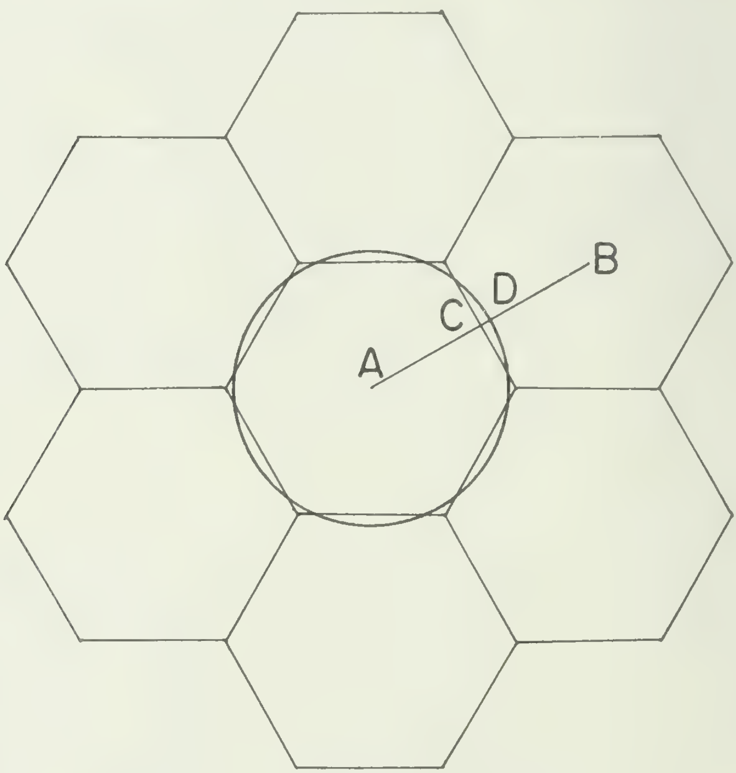


Fig. 11-1

price is greater than the price $m'_c = 0.1789b$, as the true minimum-minimorum price stands in the present case at $m'_c = 0.1789b$. It follows that only if the hexagonal competitive mill price is kept as low as this minimum price m'_c and in addition if zero profit obtains, will the market necessarily be dynamically stable. A hexagonal market with $m'_h = m'_c$ and $K'_{*h} < K'_{*c}$ but greater than the K'_{*h} of Table 10-2 is, therefore, feasible. But a final constraint, part and parcel of the possible events just described, must be added. This is done in Section III below for only after the constraint is added can the question of viability of K'_{*h}'s in the competitive space economy be fully answered.

III. *Broadening the model of spatial competition*

To fulfill present objectives, a simple model of dynamic market adjustments must be set forth in which the price differentials described above are taken as our departure point. Then finally the stable f.o.b. mill equilibrium system can be specified.

Stability conditions of the new model

For the new model the equations from Chapter 10 (10-III-1) to (10-III-3) remain, reappearing here as (11-III-1) to (11-III-3):

$$Q' = 2\rho \int_0^{\pi/\rho} \int_0^{K'_*/\cos\theta} \{b - (m' + K)\}\, K\, dK\, d\theta; \qquad (11\text{-III-}1)$$

$$Y' = m'Q' - F; \qquad (11\text{-III-}2)$$

$$K'_0(K'_*) = p' - m', \quad b > p' > m'. \qquad (11\text{-III-}3)$$

But replace (10-III-4) by the condition (11-III-4) and replace (10-III-5) by the two adjustment equations (11-III-5) and (11-III-6):

$$m' = \min m'_c = 0.1789b; \qquad (11\text{-III-}4)$$

$$\frac{dN'}{dt} = \gamma Y', \quad \gamma > 0; \qquad (11\text{-III-}5)$$

$$K'_* = \Psi(N'), \quad \Psi' < 0, \qquad (11\text{-III-}6)$$

where a new unknown N' stands for the number of firms (per square mile), and t for the time variable. Equation (11-III-5) assumes that

positive profits encourage new entry over time while negative profits force the exit of firms; in turn, $dN'/dt = 0$ if and only if $Y' = 0$. Equation (11-III-6) is a technical constraint; it establishes the relation that the greater the number of firms entering an industry, the smaller will be the trading area of a seller.

The new system to this point contains six equations in six unknowns. It appears that the system can have a stable solution. In fact, (11-III-5) plays *the role* of built-in stabilizer. To understand this effect, recall that positive profit was obtainable when $m'_c = 0.1789b$ and the seller converted from a circular market to the hexagonal market shape. But (11-III-5) signifies that positive profits generate an increase in the number of sellers; in turn, via (11-III-6), an increase in the number of sellers leads to contraction of the market size of each seller. Given the minimum mill price, this contraction implies a reduction in profit Y. Thus, another equation must be set forth:

$$Y' = \phi(K'_*), \quad \phi' > 0. \tag{11-III-7}$$

Equation (11-III-7) is not an independent new equation. In fact, (11-III-7) stems from (11-III-1), (11-III-2), and (11-III-4). It implies that so long as profits continue to be positive, the built-in stabilizer (11-III-5) remains operationally effective. We can, therefore, expect a stable zero-profit competitive equilibrium, via (11-III-5), (11-III-6), and (11-III-7).

Diagrammatic illustration of the same conclusions is possible via combination of equations (11-III-5), (11-III-6), and (11-III-7) into a single equation in one variable:

$$\frac{dN'}{dt} = \gamma Y' = \gamma\phi(K'_*) = \gamma\phi(\Psi(N')) = \gamma g(N'), \quad g' < 0 \quad \text{(since } \phi' > 0 \quad \text{and} \quad \Psi' < 0\text{)}. \tag{11-III-8}$$

The solution for this equation is presented in Fig. 11-2.

Thus far the adjustment mechanism of the system has been of central concern. However, the final stable equilibrium solution may now be determined. In this context, observe initially that though the equilibrium value of the number of sellers N' cannot be determined without precise knowledge of the exact form of (11-III-6), the remaining solutions can be determined because the equilibrium system is specifiable in a form independent of (11-III-6). This equilibrium system consists of (11-III-1) through (11-III-4), and the equilibrium condition (11-III-9):

$$Y' = 0, \tag{11-III-9}$$

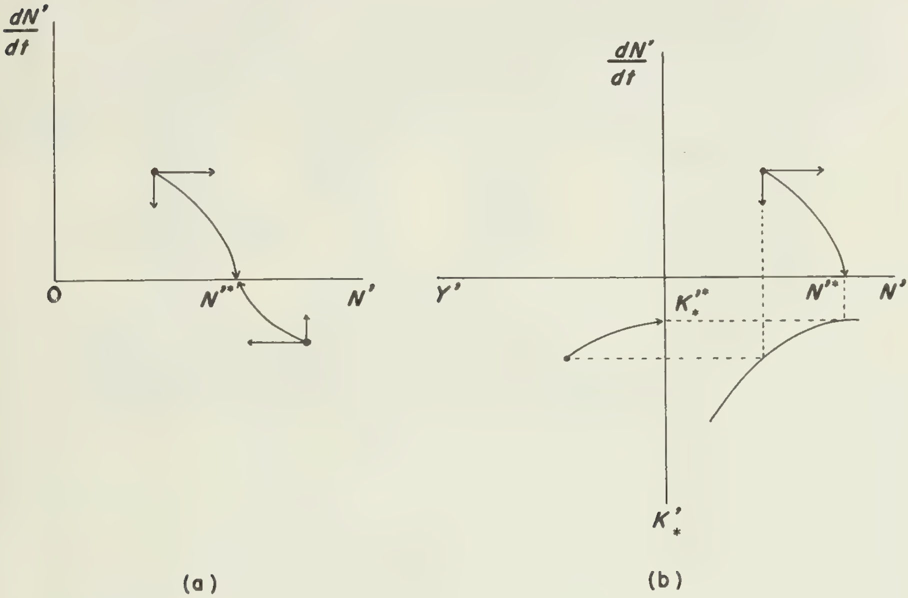


Fig. 11-2 Stability of spatial competitive equilibrium. Eq. (11-I-8) demonstrates that dN'/dt is a decreasing function of N'. Thus, as long as N' increases over time; i.e., as long as $dN'/dt > 0$, dN'/dt must decline over time. Conversely if $dN'/dt < 0$, that is if N' decreases over time, dN'/dt must rise over time. At a point where and only where $dN'/dt = 0$, the movement in N' ceases in part (a) of the figure.

The $Y'K'_*$ plane in part (b) demonstrates the function (11-I-7) and the K'_*N' plane the function (11-III-6). These being combined with part (a) illustrate the dynamic adjustment process of the system. Thus, starting from the part of positive (excess) profit due to a lack of competition, the system automatically converges to the point, $K'_*{}^*$ and N'_* reducing Y' and K'_* and increasing N' until Y' vanishes. In accordance with a movement of these variables, other variable also change over time: in particular, output Q' per firm, the market size of firms, and the maximum delivered price all decline over time, while the industry output increases over time.

where the postulation (11-III-9) is supported by the assumption of an adjustment mechanism. (11-III-9) applies in place of (11-III-5) in the long-run equilibrium. The final equilibrium thus consists of (11-III-1) through (11-III-4), (11-III-6), and (11-III-9) in six unknowns, namely Y', N', K'_*, m', Q', p'. Since the variable N' appears only in (11-III-6), it may be disregarded in solution of the other unknowns. Table 11-1 summarizes final results, attention being confined in the table to the circle and the hexagon, the only significantly relevant market area forms.

Table 11-1 *Alternative competitive equilibrium solutions compared*

	CIRCLE		HEXAGON I		HEXAGON II[a]
Market size K'	$0.4008b$		$0.3744b$		$0.3825b$
Maximum delivered price p'^*	$0.5797b$		$0.6196b$		$0.6206b$
Profit Y'^*	0.0000		0.0000		0.0000
Mill price m'^*	$0.1789b$		$0.1873b$		$0.1789b$
Output produced per firm Q'^*	$\frac{0.05b^3}{0.1789}$	$>$	$\frac{0.05b^3}{0.1873}$	$<$	$\frac{0.05b^3}{0.1789}$
The number of firms[b]	N_c		N_1		N_2

[a] The new column for Hexagon II has been obtained by making use of the same program as that employed in obtaining Table 10-2; the values for K'^* and p' have been adjusted by trial and error so that Y' approaches zero.

[b] $N_1 > N_2 > N_c$.

Theoretical implications of the spatial competition model

Consider again the questions which type of competition allows more firms to enter and which yields more outputs to society. To resolve these queries, note that except for the circle market, N' is determinable by the specification

$$N' = \frac{S}{S'_f}, \tag{11-III-10}$$

where S stands for the total area of land under study and S'_f for the market area controlled by a competitive firm. The output produced and sold by a firm, on the other hand, is given by (11-III-2), where $Y' = mQ' - F$, and (11-III-9) where $Y' = 0$, as

$$Q' = \frac{F}{m'}. \tag{11-III-11}$$

Hence, the aggregate industry output X' is given by

$$\begin{aligned} X' &= Q'N' \\ &= \frac{FS}{m'S'_f} \end{aligned} \tag{11-III-12}$$

Since F and S are both given, the aggregate industry output is determined by the mill price m', and the market area, S'_f, of a firm.

The aggregate industry outputs obtained in the two alternative com-

petitive models are fully comparable. In particular, total outputs are derived by computing the respective equilibrium mill prices and the size of the firm's market area.[3] The result of the computation shows that the product $m'S'_f$ for the original hexagon model (I) and the alternative hexagon model (II) are, respectively

$$m'_{\mathrm{I}}S'_{f\mathrm{I}} = 0.9095;$$
$$m'_{\mathrm{II}}S'_{f\mathrm{II}} = 0.9091.$$

Aggregate industry outputs are thus approximately the same for the two alternative models of hexagonal competition, with model (II) producing slightly more, viz. 0.04 percent more, than model (I).

The question as to which of the two market shapes, i.e., circle or hexagon, would yield more industry output cannot be answered by the use of equation (11-III-12), because its formula applies only to space-filling polygonal markets. However, the fact that the hexagonal markets I and II are of smaller size than the circular market's K_{*c} can be utilized. It implies that more firms can exist under hexagonal markets than under circular markets. Moreover, (11-III-11) indicates that output produced per firm under the circular market is exactly equal to that under the hexagonal market II. The aggregate industry output produced under the hexagonal market II, and hence I (which yields approximately the same industry output as II), must therefore be greater than that under the circular market. The circle is simply obliterated by competition from distant firms in economic space.

It is finally manifest that any possibility of a *kamikaze* seller entering some market can be ignored only when hexagons of size $0.3830b$ apply, with mill price m'_h and delivered price $p'_{h\mathrm{II}}$. These values alone are truly stable. The approximate equivalence in total outputs, and the slight likelihood of *kamikaze* sellers entering real world markets, prompt, however, our selecting the mill price m'_h, the delivered price $p'_{h\mathrm{I}}$, and the hexagon size K'_{*h} of model (I) as the competitive stable-equilibrium parameters. This conclusion is proposed for most sellers who practice f.o.b. mill pricing and are competing in economic space.

It might also be repeated here from Chapter 10 that we are currently engaged in empirical studies to test our conclusion with respect to the *size* of markets under the assumption that firms sell competitively over hexagonal areas. Of course, in any empirical study, rigid assumptions such as MC = 0 are relaxed, though the coefficients in Tables such as 10-2 or 11-1 continue to apply, since only the parameter b is affected, being replaced for example by the parameter b' of Chapter 10.

3. The size of the hexagonal market area S_h and circular market area S_c are given respectively by $S_h = 2\sqrt{3}K_*^2$ and $S_c = \pi K_*^2$.

IV. *The spatial competition model and price discrimination*

In this section of Chapter 11, following along the lines of [4] a model will be set forth which is a counterpart of those discussed above. In the present model, however, spatial price discrimination replaces f.o.b. mill pricing. The current objective is simply to determine the sizes and shapes of market areas which would result under discriminatory pricing practiced by either a spatial monopolist or spatial competitor. To fulfill this objective, the f.o.b. mill market area model must be reconstructed in the form of the spatial price discrimination model of Chapter 6.

Present subject matter thus requires (in fact, involves) generalization of the theory of prices and market areas previously set forth in this chapter, and Chapter 10. The generalization uncovers the full economic significance of polygon-shaped markets in a sharper manner than has thus far been achieved. Among other findings, it will be seen that higher costs of production have the same impact on mill prices and economic distances as does the number of sides of the polygon over which firms sell. This result in fact will be seen to hold true regardless of the pricing technique practiced by the firm(s). This and related findings in Section V below complete our inquiry into spatial pricing and market areas.

Aspects of a generalized discriminatory system

The fundamental meaning of a zero-profit competitive equilibrium in economic space under f.o.b. mill pricing is that the upper limit to delivered price must have been reduced to the point where preexisting windfalls are eliminated. This thesis may be appreciated best by our utilizing again the profit-maximizing delivered-price equation of the discriminatory monopolist, as derived initially in Chapter 6.[4] The equation for linear demand from Table 6-1, column p^*, is accordingly set forth as

$$p^* = \frac{b}{2} + \frac{K}{2}, \forall\, K \geqslant 0. \qquad \text{(11-IV-1)}$$

This equation (11-IV-1) is easily transformed to the mill price m^* form of (11-IV-1)′, where $m^* = p^* - K$:

$$m^* = \frac{b}{2} - \frac{K}{2}, \quad \forall\, K \ni \frac{b}{2} \geqslant K \geqslant 0. \qquad \text{(11-IV-1)}'$$

Equations (11-IV-1) and (11-IV-1)′, to repeat, relate to linear demand and an even dispersion of buyers along a line. They may be converted to that of a spatial demand function via the process set forth below.[(1)] We obtain

4. For a quick review of the original theory of discriminatory pricing over economic space, see Hoover [5], Greenhut [2], and Beckmann [1], besides Chapter 3, above.

$$q^* = \frac{b}{2} - \frac{K}{2}, \quad \forall\, K \ni \frac{b}{2} \geqslant K \geqslant 0. \tag{11-IV-2}$$

Now, equations (11-IV-1)′ and (11-IV-2) give rise to a formula for net revenue, i.e., revenue net of freight costs. Specifically, we have

$$\begin{aligned} r^* &= m^* q^* && \text{(11-IV-3)} \\ &= \left(\frac{b}{2} - \frac{K}{2}\right)\left(\frac{b}{2} - \frac{K}{2}\right), \quad \forall\, K \ni \frac{b}{2} \geqslant K \geqslant 0. \end{aligned}$$

This equation provides the net revenue of the discriminating seller on sales to any particular point K miles from his site.

The equation system (11-IV-1) to (11-IV-3) applies to an even distribution of buyers along a line. In particular, equation (11-IV-3) applies to any point on the line. Integration of (11-IV-3) provides the aggregate net revenue of a price discriminating firm selling over a market space, polygonal or circular, as indicated below:

$$\begin{aligned} R_i^* &= 2\rho \int_0^{\pi/\rho} \int_0^{K_{*i}/\cos\theta} \left(\frac{b}{2} - \frac{K}{2}\right)^2 K\, dK\, d\theta, \quad i = t, s, h && \text{(11-IV-4}_i\text{)} \\ &= \int_0^{2\pi} \int_0^{K_{*i}} \left(\frac{b}{2} - \frac{K}{2}\right)^2 K\, dK\, d\theta, \quad \text{when } i = c, \end{aligned}$$

where K_{*i}, as in preceding chapters, stands for the shortest distance from the seller's site to the market boundary and ρ stands for the number of sides of the polygon.

Now, we know that the shortest distance K_{*i} may be expressed in terms of the longest distance from the seller's location to the market boundary, K_{0i}. To recall, we have

$$K_{*t} = \frac{K_{0t}}{2}; \tag{11-IV-5$_t$}$$

$$K_{*s} = \frac{K_{0s}}{\sqrt{2}}; \tag{11-IV-5$_s$}$$

$$K_{*h} = \frac{\sqrt{3}K_{0h}}{2}; \tag{11-IV-5$_h$}$$

$$K_{*c} = K_{0c} \tag{11-IV-5$_c$}$$

In turn, the delivered prices at the maximum distances, K_{0i}, may be evaluated, for which purpose we now rewrite equation (11-IV-1) in more specific form as

$$p_{0i}{}^{*} = \frac{b}{2} + \frac{K_{0i}}{2}, \quad i = t, s, h, c. \tag{11-IV-1)''$$

The maximum distance(s) K_{0i} may then be specified via equation (11-IV-1)″ as

$$K_{0i} = 2p_{0i}{}^{*} - b, \quad i = t, s, h, c. \tag{11-IV-6$_i$}$$

Observe that $p_{0i}{}^{*}$ in (11-IV-6$_i$) stands for the profit-maximizing delivered price at the maximum distance $K = K_{0i}$.

Under competitive free entry, the maximum delivered price is parametrically given to individual firms. Thus, (11-IV-6$_i$)′ below applies to the case of competitive free entry:

$$K'_{0i} = 2p'_{0i} - b, \quad i = t, s, h, c. \tag{11-IV-6$_i$)'$$

Substituting (11-IV-6$_i$)′ into (11-IV-5$_i$) would therefore yield the parametrically given shortest distance K'_{*i}, where $K'_{*t} = p'_{0t} - b/2$, $K'_{*s} = \sqrt{2}p'_{0s} - b/\sqrt{2}$, $K'_{*h} = \sqrt{3}p'_{0h} - \sqrt{3}b/2$, and $K'_{*c} = 2p'_{0c} - b$. Substituting these expressions next into (11-IV-4$_i$) yields the net revenues applicable to the given regular polygonal or circular markets under competitive free entry *and* discriminatory pricing. The system for these markets becomes

$$R'_i{}^{*} = 2\rho \int_0^{\pi/\rho} \int_0^{K'_{*i}/\cos\theta} \left(\frac{b}{2} - \frac{K}{2}\right)^2 K \, dK d\theta, \tag{11-IV-7i}$$

for the situations where $i = t, s, h$, and

$$R'_i{}^{*} = \int_0^{2\pi} \int_0^{K'_i} \left(\frac{b}{2} - \frac{K}{2}\right)^2 K \, dK d\theta, \quad \text{for } i = c,$$

where $K'_{*i} = K_{*i}(p'_{0i})$ as specified above.[5]

Monopoly solutions

The aggregate net revenue under monopoly is given by setting the limiting value $p'_{0i} = b$ in (11-IV-7$_i$), thus removing the competitive effect.

5. Note that when $p'_i = b$ the constraint on the maximum delivered price is in effect imposed by the price intercept value b, and not by competition. On the other hand, $p'_i = b/2$ implies the disappearance of the spatial market as well as the disappearance of revenue, i.e., $K'_{*i} = 0$ and also $R'_i = 0$. Hence, p'_i must be such that $b > p'_i > b/2$ in order for the competitive spatial discriminatory model to apply.

The results relevant to a spatial monopolist who is practicing discriminatory pricing may then be obtained via integration of (11-IV-7_i). And these results for the spatial monopolist may be compared with those obtained in Chapter 10 for the case of f.o.b. mill pricing. Table 11-2 provides these comparisons.(II) As might have been expected, a discriminatory monopoly in economic space possesses a greater K_{*i}, Y_i, and Q_i than does a simple f.o.b. mill pricing spatial monopolist *regardless of the shapes of the seller's trading areas*. This conclusion holds for all market areas even though specific findings in Table 11-2 are in terms of regular polygons and circles. It is the monotonicity of the distance advantage of regular polygons over the nearest-similar irregular polygons that have identical area size which indicates the general application of our conclusion. The advantage of price discrimination over f.o.b. mill pricing for the spatial monopolist is, therefore, set forth as a general rule for all market areas.

Table 11-2 *Monopolistic equilibrium solutions under alternative pricing practices*

		TRIANGLE	SQUARE	HEXAGON	CIRCLE
Profit-maximizing K_{*i}	F.o.b.	$0.3750b$	$0.5303b$	$0.6495b$	$0.7500b$
	Discrimination	$0.5000b$	$0.7007b$	$0.8660b$	$1.0000b$
Maximum profit Y_i	F.o.b.	$0.0740b^4 - F$	$0.0961b^4 - F$	$0.1074b^4 - F$	$0.1104b^4 - F$
	Discrimination	$0.1072b^4 - F$	$0.1246b^4 - F$	$0.1304b^4 - F$	$0.1309b^4 - F$
Profit-maximizing Q_i	F.o.b.	$0.2960b^3$	$0.3844b^3$	$0.4296b^3$	$0.4416b^3$
	Discrimination	$0.3407b^3$	$0.4590b^3$	$0.5093b^3$	$0.5236b^3$

Competitive equilibrium solutions

Consider the zero-profit equilibrium condition. This condition requires equality of net revenues in (11-IV-7_i) with total production costs. (Because our model has assumed zero variable costs of production, the constraint applied here is even more elementary than it otherwise would be, since it merely requires net revenue to be equal to fixed cost F.) The equilibrium condition can thus be given as

$$R'_i{}^* = F,\ i = t, s, h, c. \qquad (11\text{-IV-}8_i)$$

Equations (11-IV-8_i), in turn, are reducible to fourth-order polynomials in p'_i. Specifically, solving (11-IV-7_i) and (11-IV-8_i) simultaneously yields

$$1.29904\,p'^{4}_{t} - 4.98855b\,p'^{3}_{t} + 6.83330b^2\,p'^{2}_{t} - 3.74140b^3\,p'_{t} + 0.70476b^4 - F = 0; \quad (11\text{-IV-}9_t)$$

$$2.66666\,p'^{4}_{s} - 9.66291b\,p'^{3}_{s} - 12.49291b^2\,p'^{2}_{s} - 6.57976b^3p'_{s} + 1.20773b^4 - F = 0; \quad (11\text{-IV-}9_s)$$

$$4.33008\,p'^{4}_{h} - 14.97840b\,p'^{3}_{h} + 18.57053b^2\,p'^{2}_{h} - 9.50177b^3\,p'_{h} + 1.70992b^4 - F = 0; \quad (11\text{-IV-}9_h)$$

$$6.28320\,p'^{4}_{c} - 20.9440b\,p'^{3}_{c} + 25.13280b^2\,p'^{2}_{c} - 12.56640b^3\,p'_{c} + 2.22530b^4 - F = 0. \quad (11\text{-IV-}9_c)$$

The relevant solution(s) for (11-IV-9_i) must assume a value or values between b and $b/2$. As shown in Table 11-3, each derived value *is unique.* Comparisons with the solutions obtained under simple f.o.b. mill pricing may also be provided, since, as in Chapter 10, we have again assumed $F = 0.05b^4$.

Table 11-3 *Competitive equilibrium solutions under alternative pricing practices*

		TRIANGLE	SQUARE	HEXAGON	CIRCLE
Maximum delivered price p'^{*}_{i}	F.o.b.	$0.8053b$	$0.6822b$	$0.6196b$	$0.5797b$
	Disc.	$0.7538b$	$0.7019b$	$0.6756b$	$0.6595b$
Market size K'^{*}_{*i}	F.o.b.	$0.2913b$	$0.3409b$	$0.3744b$	$0.4008b$
	Disc.	$0.2538b$	$0.2855b$	$0.3041b$	$0.3190b$
Quantity Q'^{*}_{i}	F.o.b.	$\frac{0.05b^3}{0.2227}$	$\frac{0.05b^3}{0.2002}$	$\frac{0.05b^3}{0.1873}$	$\frac{0.05b^3}{0.1789}$
	Disc.	$0.1283b^3$	$0.1274b^3$	$0.1260b^3$	$0.1190b^3$

V. *Alternative pricing practices under free entry*

Table 11-3 indicates that the maximum delivered prices for the zero-profit equilibrium are *in all situations except one* greater than those derived under f.o.b. pricing. In particular, Table 11-3 demonstrates that only the most inefficient regular polygon, i.e., the triangle, requires discriminatory pricing under a zero-profit competitive equilibrium[6] in

6. If competition drives the maximum delivered price p'_t down to $0.7538b$, then the f.o.b. system would yield negative profits, while the discriminatory system would yield zero profits. In fact, the f.o.b. system requires p'_t to be at least as high as $0.8053b$ in order for it to yield nonnegative profits.

economic space, provided further that $F = 0.05b^4$. In all other cases, f.o.b. mill pricing yields lower maximum delivered prices than does discriminatory pricing. Accordingly, f.o.b. mill pricing must dominate in a highly competitive situation in economic space, provided further, to repeat, that fixed costs are so relatively insignificant that market areas are comparatively small. Small market areas are just not a profitable subject for the price discrimination system, which systematically involves high mill prices to buyers located closest to the seller. These small areas do not lead to discriminatory systems everywhere in the space because a nondiscriminatory seller or sellers at a distance could so limit the sales radius of the discriminatory seller as to leave him with insufficient sales to cover his costs. In effect, he is (and all other sellers are) thus compelled to price f.o.b. mill.

A somewhat asymmetric result with respect to past findings therefore arises for market areas under discriminatory pricing. This asymmetry applies in the sense that not all market areas yield the same relations under conditions of substantial competitive entry. The alternative revenue curves depicted in Fig. 11-3 point to several of the relevant forces. The figure indicates that when fixed costs are as low as F_0F_0, e.g., $0.05b^4$, the p'_t applicable to the zero-profit equilibrium is higher for f.o.b. mill pricing than for discriminatory pricing. In contrast, p'_s, p'_h, and p'_c are all lower for f.o.b. mill pricing than they are for price discrimination. These relations, of course, also underscore Table 11-3.

Fig. 11-3 further implies that when fixed cost is as high as F_1F_1, both p'_s and p'_t are higher under f.o.b. mill pricing than they are under discriminatory pricing. Indeed, as fixed costs are raised to still higher levels, all p'_i turn out to be higher under f.o.b. mill pricing than under spatial price discrimination.

Recall now from Chapter 8 that high (fixed) cost relative to demand implies the continuance (i.e., dominance) of discriminatory pricing over f.o.b. mill pricing when buyers are distributed evenly along a line. In a similar key, deviations in shape from the circular market area have the same effect on pricing practices as do higher costs of production. The triangular market shape has been noted to be less efficient (more costly) than are other regular polygonal market shapes. Accordingly, the triangular market shape is more likely to require (or induce) discriminatory pricing than the square, the hexagon, . . . , the circle. To restate this finding in a more general way: the greater the number of sides of a market area, the lower are the firm's transport costs to buyers; in turn, the lower the costs, the smaller is the equilibrium zero-profit market area, and hence the more likely it is that f.o.b. mill pricing will prevail in the long run. Manifestly, the classical principle which limits discriminatory pricing only to the pricing practices of a simple monopoly

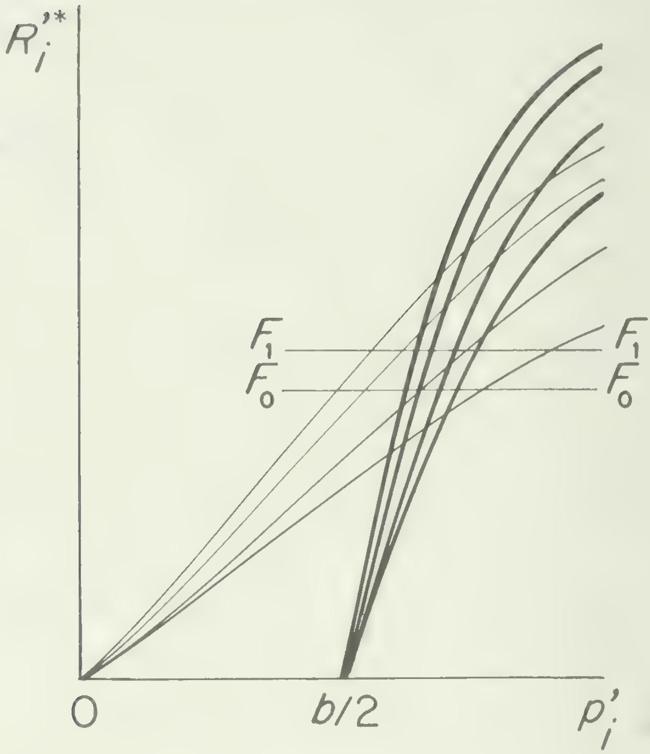


Fig. 11-3 Revenue curves for alternative market shapes under alternative pricing systems. The heavier curves, from above, stand respectively for the revenues related to the circle, hexagon, square, and triangle markets under the discriminatory system, while the lighter curves stand for the counterpart revenues under the f.o.b. system.

does not extend to the space economy, where competitors may also discriminate in price with great advantage. Our theory thus involves a set of related conclusions, namely, that spatial competitors sell over hexagon-shaped market areas, and in contrast to classical theory often may be expected to practice discriminatory pricing in the long run.

References

[1] M. Beckmann. *Location Theory.* New York, 1968.

[2] M. L. Greenhut. *Plant Location in Theory and in Practice.* Chapel Hill, N.C., 1956. Appendix to chapter 6.

[3] M. L. Greenhut, H. Ohta, and M. Proctor. "The Stability of Alternative Market Area Shapes and Sizes Under Conditions of Competition in Economic Space." Paper delivered at the Second World Congress of the Econometric Society, Cambridge, England, 1970.

[4] M. L. Greenhut, H. Ohta, and J. Scheidell. "A Model of Market Areas Under Price Discrimination." *Western Economic Journal*, 10 (Dec. 1973):402–13.
[5] E. M. Hoover. "Spatial Price Discrimination." *Review of Economic Studies*, 1937, 4:182–91.
[6] E. S. Mills and M. R. Lav. "A Model of Market Areas with Free Entry." *Journal of Political Economy*, 1964, 72:278–88.

Mathematical notes

Chapter 2

(I). The conclusion given in the text can be generalized to the case of an even distribution of buyers over a plain, a generalization which disproves the arguments of K. G. Denike and J. B. Parr [3, p. 51] and C. A. Gannon [4, pp. 295–97] that the shape of the aggregate demand curve depends on the individual demand curves. To approximate this significant relationship, consider

$$Q = \frac{\int_0^{K_0 t} 2\pi K q(m + tK)\, d(tK)}{t} \tag{a}$$

$$= \frac{2\pi}{t^2} \int_0^{q^{-1}(0)-m} x q(m + x)\, dx,$$

where q stands for the point demand density in conformance with the usual convention, i.e., q is equivalent to f^{-1} in the text divided by n; also $x = tK$ is a continuous variable, t being a small positive constant representing the unit freight rate. The first derivative of Q is

$$\frac{dQ}{dm} = \frac{2\pi}{t^2}\left[- (q^{-1}(0) - m) q(q^{-1}(0)) + \int_0^{q^{-1}(0)-m} x q'(m + x) dx\right] \tag{b}$$

$$= \frac{2\pi}{t^2}\int_0^{q^{-1}(0)-m} x q'(m + x) dx < 0 \quad \text{since } q(q^{-1}(0)) = 0 \text{ and } q' < 0.$$

The second derivative, in turn, is

$$\frac{d^2Q}{dm^2} = \frac{2\pi}{t^2}\left[- (q^{-1}(0) - m) q'(q^{-1}(0)) + \int_0^{q^{-1}(0)-m} x q''(m + x) dx\right]. \tag{c}$$

At first glance the sign of (c) does not appear to be definite, as Gannon [4, p. 297] claimed, since the first term in the bracket is nonnegative while the second term could be negative, zero, or positive. However, the

sign actually is unambiguously determinable, as (c) can easily be reduced to

$$\frac{d^2Q}{dm^2} = \frac{2\pi}{t^2}\left(-(q^{-1}(0)-m)q'(q^{-1}(0)) + [xq'(m+x) - q(m+x)]_0^{q^{-1}(0)-m}\right) \tag{c'}$$

$$= \frac{2\pi}{t^2}[-q(q^{-1}(0)) + q(m)]$$

$$= \frac{2\pi}{t^2} q(m) > 0, \quad \forall\, m < q^{-1}(0).$$

Thus, the monopoly spatial demand function is *always* concave from above regardless of the shape (concavity or convexity) of the individual gross demand function [6].

II. Ming J. Hwang proposes the following proof of the location of the average buyer along a line market. Let the total number of buying points be n; total distance T may then be given as

$$T = \frac{(1+2+\cdots+n)K_0}{n}. \tag{a}$$

The average buyer is located at the halfway market point:

$$\frac{T}{n} = \frac{(n+1)K_0}{2n}, \therefore \lim_{n\to\infty}\left(\frac{T}{n}\right) = \frac{K_0}{2}. \tag{b}$$

If all buyers are distributed continuously along the line such that their density is one buyer per unit of distance, total density is specifiable as

$$\int_0^{K_0} dK = K_0, \tag{c}$$

and both total and average distance are, respectively

$$\int_0^{K_0} KdK = \frac{K_0^2}{2}, \quad \text{and} \quad \frac{\left(\frac{K_0^2}{2}\right)}{K_0} = \frac{K_0}{2}. \tag{d}$$

(III). The spaceless demand function is $Q = (f(0) - p)/a$. For comparative purposes, let mill price m = delivered price $p = 0$. The quantity demanded in the spaceless economy is then $Q = f(0)/a$ since all buyers have the same demand function $q = f(0)/na$. In the case of economic distances, the buyer at the seller's location who is charged mill price $m = 0$ pur-

chases $q = f(0)/na$. The second buyer, who is assumed to be one unit away from the seller's location, pays the delivered price $m + tK = 0 + t = t$ so that the quantity demanded by him is slightly less than $q = f(0)/na$. The last buyer, being n units of distance away from the seller's location, pays $m + tK = 0 + tK_0 = f(0)$, so that the quantity he demands is $q = (f(0) - p)/na = (f(0) - f(0))/na = 0$. When $m = 0$, buyers located along a line are on the average subject to a delivered price $f(0)/2$ higher than the spaceless economy price, and the total quantity demanded is therefore $\frac{1}{2}$ less than it is in the counterpart spaceless economy.

(IV). Equating spatial marginal revenue with marginal production cost might be detailed with advantage. Since we must equate in the line distribution case $c = f(0) - \frac{3}{2}(2af(0)Q)^{1/2}$, then

$$\frac{3}{2}\sqrt{(2af(0)Q)} = f(0) - c, \tag{a}$$

$$\frac{9}{4}(2af(0)Q) = (f(0) - c)^2, \quad \text{and} \tag{b}$$

$$Q = \frac{(2/9)(f(0) - c)^2}{af(0)}. \tag{c}$$

Correspondingly in the case of buyers distributed over a plain in the density $1/\pi f(0)^2$, we must equate $c = f(0) - \frac{4}{3}(3af(0)^2Q)^{1/3}$. This yields

$$\frac{4}{3}(3af(0)^2Q)^{1/3} = f(0) - c; \tag{a$'$}$$

$$\frac{64}{27}(3af(0)^2Q) = (f(0) - c)^3; \tag{b$'$}$$

$$Q = \frac{(9/64)(f(0) - c)^3}{af(0)^2}. \tag{c$'$}$$

(V). Substituting the equilibrium Q of (2-17) in $p = f(0) - aQ$ yields the equilibrium nonspatial price:

$$\begin{aligned} p &= f(0) - a\left(\frac{f(0) - c}{2a}\right) \qquad (1) \\ &= \frac{2f(0) - f(0) + c}{2} \\ &= \frac{f(0)}{2} + \frac{c}{2}. \end{aligned}$$

For the spatial line case, since $m = f(0) - (2af(0)Q)^{1/2}$, a similar substitution from (2-18) yields the equilibrium mill price:

$$
\begin{aligned}
m &= f(0) - \left[\frac{2af(0)(2/9)(f(0)-c)^2}{af(0)}\right]^{1/2} \qquad (2)\\
&= f(0) - \left[\frac{4}{9}(f(0)-c)^2\right]^{1/2}\\
&= \frac{3f(0) - 2f(0) + 2c}{3}\\
&= \frac{f(0)}{3} + \frac{2c}{3}.
\end{aligned}
$$

For the spatial plain case, since $m = f(0) - (3a - f(0)^2 Q)^{1/3}$, substitution from (2-19) yields

$$
\begin{aligned}
m &= f(0) - \left[\frac{3af(0)^2(9/64)(f(0)-c)^3}{af(0)^2}\right]^{1/3} \qquad (3)\\
&= f(0) - \left[\frac{27}{64}(f(0)-c)^3\right]^{1/3}\\
&= f(0) - \left[\frac{3}{4}(f(0)-c)\right]\\
&= \frac{f(0)}{4} + \frac{3c}{4}.
\end{aligned}
$$

Chapter 4

(I). Substitute the inverse of (4-3), i.e., (4-2), into (4-3D). This process yields

$$\frac{2n(b-m)}{3b} + 1 \geqslant i > 2n(b-m) - \frac{1}{3}. \qquad \text{(a)}$$

Because i is an integer and the interval of i is greater than unity, i.e., $\frac{4}{3}$, inequality (a) provides either one or two values for i, depending on the value of m. Thus, i in general is not a function of m. This result, of course, stems from the fact that SMR is not a one-to-one function. Nevertheless (a) can be rewritten in the more revealing form (a)′:

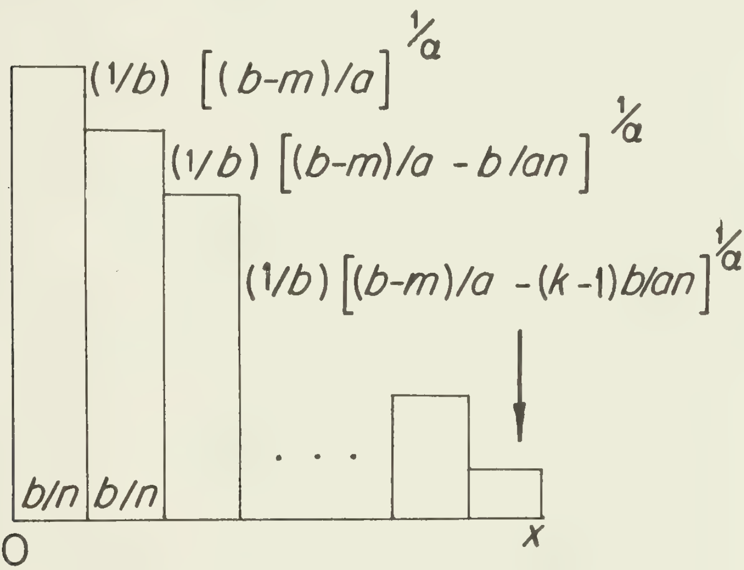
$$\frac{2(b-m)}{3b}+\frac{1}{n}\geqslant\frac{i}{n}>\frac{2(b-m)}{3b}-\frac{1}{3n} \tag{a}'$$

And we see that if n is a large number, i/n may be approximated by $2(b-m)/3b$.

(II).

$$\begin{aligned} Q &= \frac{1}{n}\sum_{i=1}^{k}\left[\frac{b-m}{a}-\frac{(i-1)b}{an}\right]^{1/\alpha} \\ &= \frac{1}{b}\sum_{i=1}^{k}\frac{b}{n}\left[\frac{b-m-(i-1)b/n}{a}\right]^{1/\alpha}, \end{aligned}$$

where $(1/b)[(b-m)/a]^{1/\alpha}$ is the demand value over the first distance interval b/n, $(1/b)[(b-m)/a-b/an]^{1/\alpha}$ applies to the market in the next distance interval b/n, and $(1/b)[(b-m)/a-(k-1)b/an]^{1/\alpha}$ applies to the farthest market in the seller's market area, as assumed in the figure.



Let $x=(i-1)b/n$. Then

$$Q=\frac{1}{b}\int_{0}^{(k-1)b/n}\left(\frac{b-m-x}{a}\right)^{1/\alpha}dx,$$

where the upper limit for x is given by substituting $i=k$ into $x=(i-1)b/n$.

(III). Equation (4-7)″ may be rearranged as follows:

$$Q^* = \frac{a\alpha}{b(1+\alpha)}\left(\frac{b-m}{a}\right)^{(1+\alpha)/\alpha};$$

$$\frac{b(1+\alpha)}{a\alpha}Q^* = \left(\frac{b-m}{a}\right)^{(1+\alpha)/\alpha};$$

$$\frac{b-m}{a} = \left[\frac{b(1+\alpha)}{a\alpha}Q^*\right]^{\alpha/(1+\alpha)};$$

$$\therefore m = b - a\left[\frac{b(1+\alpha)}{a\alpha}Q^*\right]^{\alpha/(1+\alpha)};$$

$$\therefore mQ^* = bQ^* - a\left[\frac{b(1+\alpha)}{a\alpha}\right]^{\alpha/(1+\alpha)} Q^{*(1+2\alpha)/(1+\alpha)}.$$

Then $d(mQ^*)/dQ^*$ gives SMR, referred to below as R, and its inverse, as in (4-8):

$$R = b - a\left(\frac{1+2\alpha}{1+\alpha}\right)\left[\frac{b(1+\alpha)}{a\alpha}Q^*\right]^{\alpha/(1+\alpha)};$$

$$b - R = a\left(\frac{1+2\alpha}{1+\alpha}\right)\left[\frac{b(1+\alpha)}{a\alpha}\right]^{\alpha/(1+\alpha)} Q^{*\alpha/(1+\alpha)}.$$

$$\therefore Q^{*\alpha/(1+\alpha)} = \frac{1}{a}\left(\frac{1+\alpha}{1+2\alpha}\right)\left[\frac{a\alpha}{b(1+\alpha)}\right]^{\alpha/(1+\alpha)}(b-R), \quad \text{and}$$

$$Q^* = \left(\frac{a\alpha}{b(1+\alpha)}\right)\left(\frac{1+\alpha}{1+2\alpha}\right)^{(1+\alpha)/\alpha}\left(\frac{b-R}{a}\right)^{(1+\alpha)/\alpha}$$

(IV). In (4-6)′, individual demand was shown to be

$$q = \left(\frac{1}{n}\right)\left[\frac{(b-m)}{a} - \frac{(i-1)b}{an}\right]^{1/\alpha};$$

$$nq^\alpha = \left[\frac{(b-m)}{a} - \frac{(i-1)b}{an}\right];$$

$$m = b - a(nq)^\alpha - \frac{(i-1)b}{n};$$

$$mq = bq - aq(nq)^\alpha - \frac{(i-1)bq}{n}.$$

R in (4-9) is derivable by differentiating this revenue function with respect to q:

$$R = \frac{d(mq)}{dq} = b - [a(nq)^{\alpha} + a\alpha(nq)^{\alpha-1}(nq)] - (i-1)b/n$$
$$= b - a\,(\alpha+1)nq^{\alpha} - (i-1)b/n$$

(V).

$$Q = \frac{1}{n}\sum_{i=1}^{k}\left[\frac{b-R-(i-1)b/n}{a(\alpha+1)}\right]^{1/\alpha}$$
$$= \frac{1}{b}\sum_{i=1}^{k}\frac{b}{n}\left[\frac{b-R-(i-1)b/n}{a(\alpha+1)}\right]^{1/\alpha}$$
$$\therefore \lim_{n\to\alpha} Q = \frac{1}{b}\int_0^{(k-1)b/n}\left[\frac{b-R-x}{a(\alpha+1)}\right]^{1/\alpha} dx.$$

Substituting R for m in the formula $k = (\frac{1}{2}) + (b-m)n/b$ and taking the limit gives

$$\lim_{n\to\infty} Q = \frac{1}{b}\int_0^{b-R}\left[\frac{b-R-x}{a(\alpha+1)}\right]^{1/\alpha} dx.$$

(VI). From (4-10)′, we obtain

$$\lim_{n\to\infty} Q = \frac{1}{b}\left[(-a\alpha)\left(\frac{b-R-x}{a(\alpha+1)}\right)^{(1+\alpha)/\alpha}\right]_0^{b-R}$$
$$= \frac{1}{b}\,a\alpha\left(\frac{b-R}{a(\alpha+1)}\right)^{(1+\alpha)/\alpha}$$
$$= \frac{a\alpha}{b}\left(\frac{1}{1+\alpha}\right)^{(1+\alpha)/\alpha}\left(\frac{b-R}{a}\right)^{(1+\alpha)/\alpha}$$

Chapter 5

(I).

$$p\left(\frac{e-1}{e}\right) - K = 0;$$

$$\frac{dp}{dK}\left(\frac{e-1}{e}\right) + p\left(\frac{e\dfrac{de}{dK} - e\dfrac{de}{dK} + \dfrac{de}{dK}}{e^2}\right) = 1;$$

$$\frac{dp}{dK}\left(\frac{e-1}{e}\right)+p\left(\frac{\frac{de}{dp}\frac{dp}{dK}}{e^2}\right)=1;$$

$$\frac{dp}{dK}\left[\left(\frac{e-1}{e}\right)+\frac{de}{dp}\frac{p}{e^2}\right]=1;$$

$$\frac{dp}{dK}\left[\frac{e-1}{e}+\frac{\epsilon}{e}\right]=1, \text{ where } \epsilon=\frac{de}{dp}\frac{p}{e}.$$

$$\therefore \frac{dp}{dK}=\frac{e}{\epsilon-(1-e)}, \quad \text{provided} \quad \frac{\epsilon-(1-e)}{e}\neq 0.$$

(II). Define e as $e=-(dq/dp)(p/q)$ so that $dq/dp=-qe/p$.

The second derivative of q with respect to p is

$$\frac{d^2q}{dp^2}=-\left[\frac{de}{dp}\frac{q}{p}+\frac{(1+e)(dq/dp)}{p}\right].$$

Substituting this into (5-5) yields

$$\frac{(dq/dp)^2}{q}+\left[\frac{de}{dp}\frac{q}{p}+\frac{(1+e)(dq/dp)}{p}\right]<0$$

$$\Leftrightarrow \frac{de}{dp}\frac{q}{p}<-\frac{dq}{dp}\left(\frac{dq}{dp}\frac{1}{q}+\frac{1+e}{p}\right)=\frac{eq}{p^2}\Leftrightarrow \epsilon=\frac{de}{dp}\frac{p}{e}<1.$$

Chapter 6

(I).

$$pq=\left(-\sqrt{(2bq)}+\frac{3}{2}b\right)q. \tag{1a}$$

$$\frac{d(pq)}{dq}=\frac{3}{2}\left[b-\sqrt{(2bq)}\right]. \tag{2a}$$

Setting (2a) equal to K yields

$$q=\frac{1}{2b}\left(b-\frac{2}{3}K\right)^2, \quad \frac{3}{2}b\geqslant K\geqslant 0. \tag{3a}$$

Profit-maximizing sales distance $K_0{}^*$ is then given by

$$\frac{dR}{dK}=\frac{dR}{dq}\frac{dq}{dK}=-\frac{dR}{dq}\left(\frac{2}{3b}\right)\left(b-\frac{2}{3}K\right)=0, \tag{4a}$$

where R stands for integrated revenue, i.e., $\int_0^{K_0} pq\, dK$.

$$\therefore K_0{}^*=\frac{3}{2}b. \tag{5a}$$

Observe from (5a) and (3a) that point demand q becomes zero at the profit-maximizing sales distance $K_0{}^*=\frac{3}{2}b$.

(II). Substituting (3a) into the revenue function (1a), and integrating it with respect to market size K_0 yields the profit-maximizing revenue R^*. Thus:

$$\begin{aligned} R^* &= \int_0^{3b/2} (pq)dK \\ &= \frac{1}{2b}\int_0^{3b/2}\left(\frac{3}{2}b-(b-2K/3)\right)\left(b-\frac{2}{3}K\right)^2 dK=\frac{3}{16}b^3. \end{aligned} \tag{6a}$$

(III). The total maximum profit (or net revenue) can be obtained as follows:

$$\Pi^*=\int_0^{3b/2}(pq-Kq)dK=\frac{1}{4b}\int_0^{3b/2}\left(b-\frac{2}{3}K\right)^3 dK=\frac{3}{32}b^3. \tag{6a$'$}$$

(IV). Profit-maximizing output Q^* is:

$$Q^*=\int_0^{3b/2} q\, dK=\int_0^{3b/2}\frac{1}{2}b\left(b-\frac{2}{3}K\right)^2 dK=\frac{1}{4}b^2. \tag{7a}$$

(V). Substituting (3a) into the demand function (a) provides the delivered price:

$$p^*=\frac{2}{3}K+\frac{1}{2}b, \tag{8a}$$

with the constraint $(\frac{3}{2})b \geqslant K \geqslant 0$ given by equation (5a). Equation (8a) later is shown in Fig. 6-2 as the price line $p_{\mathrm{I}}p_{\mathrm{I}}$. Profit-maximizing mill prices then are determined by way of definition as

$$m^*=p^*-K=\frac{1}{2}b-\frac{1}{3}K. \tag{8a$'$}$$

(VI). We find

$$m + K = -\sqrt{(2bq)} + \frac{3}{2}b. \tag{9a}$$

$$\therefore q = \frac{1}{2b}\left[\frac{3}{2}b - (m + K)\right]^2. \tag{10a}$$

Hence, the aggregate output Q (not Q^*) that is produced by the seller for whom sales radius K_0 applies, and the related revenue are

$$\begin{aligned} Q = \int_0^{K_0} q\, dK &= -\frac{1}{6b}\left[\frac{3}{2}b - (m + K)\right]^3 \Big/_0^{K_0} \\ &= \frac{-1}{6b}\left(\frac{3}{2}b - m - K_0\right)^3 + \frac{1}{6b}\left(\frac{3}{2}b - m\right)^3, \end{aligned} \tag{11a}$$

and

$$R = mQ = \frac{m}{6b}\left[-\left(\frac{3}{2}b - m - K_0\right)^3 + \left(\frac{3}{2}b - m\right)^3\right]. \tag{12a}$$

In order to obtain the profit-maximizing set $\{m^*, K_0^*\}$, we form the following equations:

$$\frac{\partial R}{\partial m} = \frac{1}{6b}\left[9m^2 - (18b - 6K_0)\, m + K_0^2 - \frac{9}{2}bK_0 + \frac{27}{4}b^2\right] = 0; \tag{13a}$$

$$\frac{\partial R}{\partial K_0} = \frac{m}{2b}\left(\frac{3}{2}b - m - K_0\right)^2 = 0. \tag{14a}$$

Since equations (13a) and (14a) contain two unknowns, namely, m and K_0, unique solutions are obtainable:

$$m^* = \frac{3}{8}b,\ K_0^* = \frac{9}{8}b. \tag{15a}$$

Other solutions may be found as well, but they are not profit-maximizing.

(VII). Substituting now $\{m^*, K^*\}$ into (11a) and (12a) yields

$$Q^* = \left(\frac{3}{4}\right)^5 b^2, \quad \text{and} \tag{16a}$$

$$\Pi^* = \frac{1}{2}\left(\frac{3}{4}\right)^6 b^3. \tag{17a}$$

(VIII). Substituting m^* into equation (5) in Model II and combining it with equation (6) of Model II provides the set of profit-maximizing delivered prices in terms of distance K:

$$p^* = m^* + K = \frac{3}{8}b + K, \tag{18a}$$

where the constraint $\frac{9}{8}b > K > 0$ is given by (15a).

(IX). For a proof, let $p = f(q)$ as before stand for a demand function of the general form. Then profit maximization under MC = 0 requires, at every K, the condition $\text{MR} = (d/dq)\ [p - K]q = f'(q)q + [f(q) - K] = 0$. Differentiating this condition with respect to K yields

$$f'' \frac{dq}{dp}\frac{dp}{dK}q + f' \frac{dq}{dp}\frac{dp}{dK} + f' \frac{dq}{dp}\frac{dp}{dK} - 1 = 0.$$

$$\therefore 1 - \frac{dp}{dK} = 1 - \frac{1}{2 + (f''/f')q}.$$

Clearly, the rate of freight absorption depends upon the convexity of the demand, i.e., f''. Since $q > 0$ and $f' < 0$, concave demand, i.e., $f'' > 0$, requires the absorption rate to be less than $\frac{1}{2}$, linear demand requires it to be equal to $\frac{1}{2}$, and convex demand requires it to be greater than $\frac{1}{2}$.

(X). Proof of this relationship between concavity and the variable x may be set forth easily. When the demand is

$$p = -\alpha q^x + \beta, \tag{i}$$

the slope is given by

$$\frac{dp}{dq} = -\alpha x q^{x-1}, \tag{ii}$$

and the concavity is measured as

$$\frac{d^2p}{dq^2} = -\alpha x(x-1)q^{x-2}. \tag{iii}$$

Manifestly the slope is negative when $\alpha, x > 0$, and the concavity positive, zero, or negative when $(x - 1) \lesseqgtr 0$ (since α, x, and q^{x-2} are all positive). Thus the demand curve is concave, linear, or convex, respectively, when x is less than, equal to, or greater than unity.

Chapter 7

(I). Observe from (7-6)″ and the requirement $p_f \geqslant 0$ that

$$p_f = \frac{a_1 b_0 - a_0 b_1}{2b_0} + \frac{b_1}{2b_0} p_r \geqslant 0.$$

By substituting for p_r from (7-6)′, we obtain

$$p_f = \frac{a_1 b_0 - a_0 b_1}{2b_0} + \frac{b_1}{2b_0}\left[\frac{a_0 c_1 - a_1 c_0}{2c_1} + \frac{c_0}{2c_1} p_f\right] \geqslant 0.$$

Then

$$p_f = \frac{2a_1 b_0 c_1 - a_0 b_1 c_1 - a_1 b_1 c_0}{4b_0 c_1 - b_1 c_0} = \frac{\beta}{\alpha} \geqslant 0.$$

Moreover, as noted above, Fig. 7-2 requires that at least α or β be positive; hence $\alpha > 0$, $\beta \geqslant 0$. But for intersection to occur in the first quadrant, p_r in (7-6)″ must also be $\geqslant 0$; i.e.,

$$p_r = \frac{a_0 b_1 - a_1 b_0}{b_1} + \frac{2b_0}{b_1} p_f \geqslant 0 \quad \text{or} \quad a_0 b_1 - a_1 b_0 + 2b_0 \frac{\beta}{\alpha} \geqslant 0.$$

(II). Suppose the following demands characterize the market:

$$p_r = a_0 - b_0 q_r^{1/2} - c_0 \Sigma q_s; \tag{i}$$

$$p_f = a_1 - b_1 q_r - c_1 \Sigma q_s^{1/2}. \tag{ii}$$

Reaction functions are

$$q_r = \left(\frac{2}{3}\frac{a_0}{b_0} - \frac{2}{3}\frac{c_0}{b_0}\Sigma q_s\right)^2; \tag{iii}$$

$$q_r = \frac{a_1}{b_1} - \frac{3}{2b_1}(\Sigma q_s)^2. \tag{iv}$$

And

$$\frac{dq_r}{d\Sigma q_s} = 2\left(\frac{2}{3}\frac{a_0}{b_0} - \frac{2}{3}\frac{c_0}{b_0}\Sigma q_s\right)\left(-\frac{2}{3}\frac{c_0}{b_0}\right) < 0; \tag{v}$$

$$\frac{d^2q_r}{d(\Sigma q_s)^2} = \left(\frac{8}{9}\frac{c_0}{b_0}\right)^2 > 0. \tag{v)$'$}$$

While

$$\frac{dq_r}{d\Sigma q_s} = -\frac{3}{4}\frac{1}{b_1}(\Sigma q_s)^{-(1/2)} < 0; \tag{vi}$$

$$\frac{d^2q_r}{d(\Sigma q_s)^2} = \frac{3}{8}\frac{1}{b_1}(\Sigma q_s)^{-(3/2)} > 0. \tag{vii}$$

They are all concave upwards, etc.

Chapter 8

(I). Equation (8-22) by substitution of (8-19) can be rewritten as

$$\Pi_i = \sum_{j=1}^{n} f_j(q_{\cdot j})q_{ij} - C_i - \sum_{j=1}^{n} T_{ij}q_{ij}, \quad \forall\ i, j. \tag{8-22$'$}$$

The first-order conditions for profit maximization are accordingly

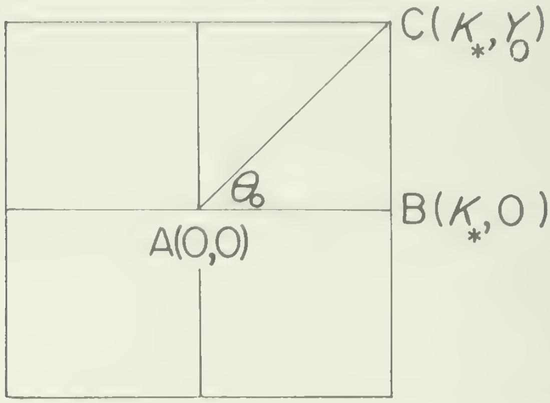
$$\frac{\partial \Pi_i}{\partial q_{ij}} = f_j(q_{\cdot j}) + f'_j(q_{\cdot j})\frac{\partial q_{\cdot j}}{\partial q_{ij}}q_{ij} - C'_i\frac{\partial q_i}{\partial q_{ij}} - T_{ij} = 0, \quad \forall\ i, j. \tag{8-22$''$}$$

where $(\partial q_{\cdot j}/\partial q_{ij})$ turns out to be unity under each firm's anticipation that its rival firm's supply, as given on the right-hand side of (8-20), is invariant. By elementary manipulation, (8-22)″ is reduced to (8-23) in the text:

$$p_j\left(1 - \frac{1}{\frac{\partial q_{\cdot j}}{\partial p_j}\frac{p_j}{q_{\cdot j}}}\frac{q_{ij}}{q_{\cdot j}}\right) - C'_i - T_{ij} = 0, \ \forall\ i, j.$$

Chapter 9

(I). The figure represents a square market area; thus $\rho = 4$. We are integrating over the area ABC and multiplying the result by $2\rho = 8$ to obtain total sales over the entire square, as there are 8 triangles of size



$\theta_0 = \pi/4$ radian; AB $= K'_*$, shortest distance from A to the perimeter.

ABC in the square. Using polar coordinates, let K vary from the origin A $= (0, 0)$ to some point on the line BC, that point depending on the angle θ. Angle θ is the relation between the distance K_* and a given coordinate on BC: it is defined as $\cos\theta = K_*/K$, where $K = \sqrt{(K_*^2 + y^2)}$, with K_* and y being a coordinate on BC, i.e., $(K, y) = (K_*, y)$; and $(K, y) = (0, 0)$ at A. For example, if $\theta = \pi/4$, $K_* =$ AB, and $K = K_0 =$ AC, $\cos \pi 4 =$ AB/AC by definition. $K = K_*/\cos\theta$ therefore takes its upper limit value when θ takes its upper limit value, i.e. $\pi/4$.

(II). Equations (9-4i) are of the form

$$Q = 2\rho \int_0^{\pi/\rho} \left\{ \frac{(b-m)K^2}{2} - \frac{K^3}{3} \Big/_0^{K_*/\cos\theta} \right\} d\theta$$

$$= 2\rho \int_0^{\pi/\rho} \left\{ \frac{(b-m)K_*^2}{2\cos^2\theta} - \frac{K_*^3}{3\cos^3\theta} \right\} d\theta.$$

From a table of integrals we know

$$\frac{\int d\theta}{\cos\theta} = \ln(\tan\theta + \sec\theta), \frac{\int d\theta}{\cos^2\theta} = \tan\theta,$$

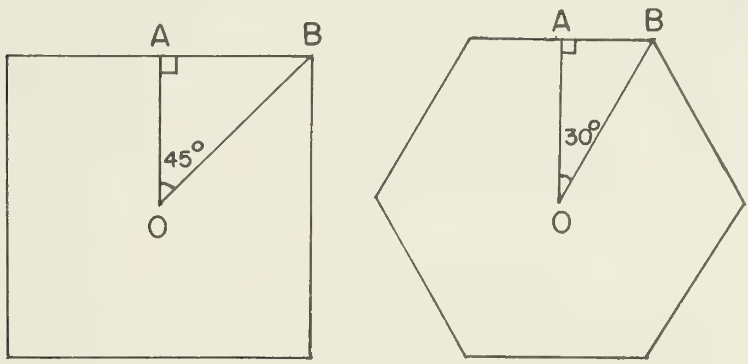
$$\text{and } \frac{\int d\theta}{\cos^n\theta} = \frac{1}{n-1}\left(\frac{\sin\theta}{\cos^2\theta}\right) + \frac{n-2}{n-1} \int \frac{d\theta}{\cos^{n-2}\theta}.$$

Thus:

$$\frac{\int d\theta}{\cos^3\theta} = \left(\frac{1}{2}\right)\left(\frac{\sin\theta}{\cos^2\theta} + \int \frac{d\theta}{\cos\theta}\right) = \left(\frac{1}{2}\right)\left[\frac{\sin\theta}{\cos^2\theta} + \ln(\tan\theta + \sec\theta)\right].$$

The θ's attain different values according to the polygon. Then, upon substitution of appropriate numerical values for the trigonometric relations, we obtain equations (9-4i) shown in the text.

(III). For a regular square market, the Pythagorean theorem indicates that $(OA)^2 + (AB)^2 = (OB)^2$. Since $OA = AB =$ the shortest distance (K_*) from the center and $OB =$ the longest distance (K_0) from the center, it follows that the longest distance $OB = \sqrt{2}$ times the shortest distance OA; i.e., $K_0 = \sqrt{2}K_*$ – see Fig. (a). On the other hand, for a hexagon market, again $(OA)^2 + (AB)^2 = (OB)^2$. Thus, from Fig. (b):



$$OB = 2(AB) \text{ or } AB = \frac{(OB)}{2}; \qquad \text{(i)}$$

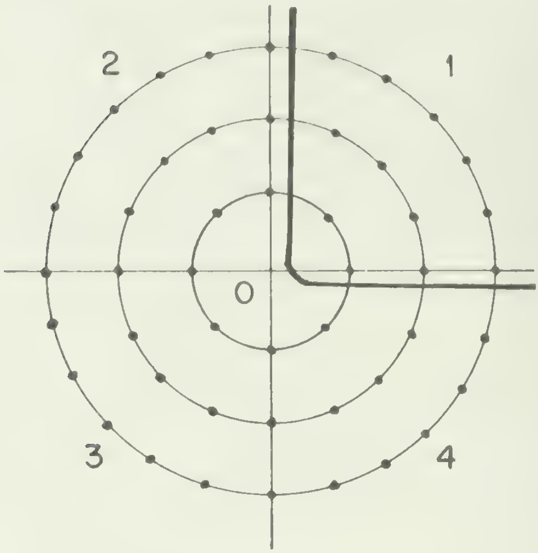
$$(OA)^2 + \left[\frac{(OB)}{2}\right]^2 = (OB)^2; \qquad \text{(ii)}$$

$$OB = \frac{2(OA)}{\sqrt{3}}; \qquad \text{(iii)}$$

$$K_0 = \frac{2K_*}{\sqrt{3}}. \qquad \text{(iv)}$$

Chapter 10

(I). Ming J. Hwang proposes the following proof of the location of the average buyer in a circular market. He utilizes the accompanying figure.



Assume the seller is located at the origin O and buyers are evenly located around the origin as in the sketch. This distribution is derived from the square form distribution used previously (pp. 180–81). Assuming every point on any circle is of equal distance from the origin, the accompanying diagram obtains. Manifestly, equal distance in economic space means equal straight-line distance to the origin. The points on the first circle have equal distance K_0/n, the points on the second circle are $2K_0/n$, etc.

In the first section there are $2 + 4 + 6 + \cdots + 2n = 2\sum_{i=1}^{n} i = n(n+1)$ points.

The total points in the circle therefore are $4n(n+1)$. And the total distances in the first section are

$$2\left(\frac{K_0}{n}\right) + 4\left(\frac{2K_0}{n}\right) + 6\left(\frac{3K_0}{n}\right) + \cdots + 2n\left(\frac{nK_0}{n}\right) = \frac{(2\sum_{i=1}^{n} i^2)K_0}{n}, \qquad \text{(i)}$$

the total distances being

$$\frac{4(2\sum_{i=1}^{n} i^2)K_0}{n}. \qquad \text{(ii)}$$

Therefore the average distance is

$$d = \frac{4(2\Sigma_{i=1}^{n} i^2)K_0/n}{4n(n+1)} = \frac{2K_0 n(n+1)(2n+1)/6}{n^2(n+1)} \tag{iii}$$

$$= \frac{2K_0}{3} + \frac{K_0}{3n}.$$

$$\therefore \lim_{n\to\infty} d = \frac{2K_0}{3}.$$

Chapter 11

(I). From the assumed demand curve, i.e.,

$$p = b - q, \tag{i}$$

and the definition of mill price, i.e.,

$$m = p - K, \tag{ii}$$

we obtain

$$p = m + K = b - q; \tag{iii}$$
$$m = b - q - K. \tag{iii)$'$}$$

Relating this m to (11-IV-1)′ yields:

$$b - q^* - K = \frac{(b-K)}{2}; \tag{iv}$$

$$q^* = \frac{(b-K)}{2}, \quad \forall\, K \ni \frac{b}{2} \geqslant K \geqslant 0. \tag{iv)$'$}$$

(II). The Q_i column for the discriminatory case is obtainable from

$$Q_i = 2\rho \int_0^{\pi/\rho} \int_0^{K_{*i}/\cos\theta} \left(\frac{b}{2} - \frac{K}{2}\right) K \, dK d\theta, \quad i = t, s, h,$$

$$= \int_0^{2\pi} \int_0^{K_{*i}} \left(\frac{b}{2} - \frac{K}{2}\right) K \, dK d\theta, \quad i = c.$$

Index

Name

Subject